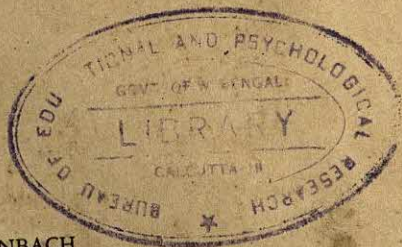


THE AMERICAN JOURNAL OF PSYCHOLOGY

165



EDITED BY
KARL M. DALLENBACH
University of Texas
AND

M. E. BITTERMAN
Bryn Mawr College

E. B. NEWMAN
Harvard University

Bureau of Ednl. & Psyl. Research

(S. C. E. R. T.)

Date 3.3.81

5651

WITH THE COÖPERATION OF

Acc...
E. G. BORING, Harvard University; W. K. ESTES, Indiana University; J. P. GUILFORD, University of Southern California; HARRY HELLSON, University of Texas; E. R. HILGARD, Stanford University; FRANCIS W. IRWIN, University of Pennsylvania; G. L. KREEZER, Washington University; D. G. MARQUIS, University of Michigan; GEORGE A. MILLER, Harvard University; R. M. OGDEN, Cornell University; W. B. PILLSBURY, University of Michigan; LEO POSTMAN, University of California; W. C. H. PRENTICE, Swarthmore College; T. A. RYAN, Cornell University.

Vol. LXX 70
MEZES HALL, THE UNIVERSITY OF TEXAS
AUSTIN, TEXAS

1957

Copyright 1957 by Karl M. Dallenbach

47.50

TABLE OF CONTENTS

ARTICLES AND NOTES

ADAMS, J. K., A Confidence Scale Defined in Terms of Expected Percentages	432
ADAMS, P. A., and POSTMAN, L., On Recent Studies of the Law of Effect in Incidental Learning	642
ANDERSON, L., and OSGOOD, C. E., Certain Relations among Experienced Contingencies, Associative Structure, and Contingencies in Encoded Messages ..	411
ATTNEAVE, F., An Inter-Ocular Color-Effect	318
BAKER, K. E., and DUDEK, F. J., Scaling Line-Lengths with a Modification of the Constant-Sum Method	81
BAKER, K. E., and DUDEK, F. J., On the Validity of the Point-Assignment Procedure in the Constant-Sum Method	268
BATTIG, W. F., BROGDEN, W. J., NAGEL, E. H., and VOSS, J. F., Transfer and Retention of Bidimensional Compensatory Tracking after Extended Practice	75
BENDIG, A. W., and VAUGHAN, C. J., Manifest Anxiety, Discrimination, and Transposition	286
BERNARDIN, A. C., and GRUBER, H. E., An Auditory Autokinetic Effect	133
BILODEAU, E. A., The Relationship between a Relatively Complex Motor Skill and Its Components	49
BILODEAU, E. A., Patterns of Internal Consistency in Multipart Skilled Performances	550
BITTERMAN, M. E., DALLENBACH, K. M., and NEWMAN, E. B., Editorial Announcements	135
BITTERMAN, M. E., and WODINSKY, J., Discrimination-Reversal in the Fish	569
BRANCA, A. A., Semantic Generalization at the Level of the Conditioning Experiment	541
BRAUN, H. W., HESSERT, R. B., and ROCKLYN, E. H., Calibrated Materials for Verbal Learning in Middle- and Old-Aged Subjects	628
BRENNAN, G., and KARLIN, L., Memory for Visual Figures by the Method of Identical Stimuli	248
BROGDEN, W. J., BATTIG, W. F., NAGEL, E. H., and VOSS, J. F., Transfer and Retention of Bidimensional Compensatory Tracking after Extended Practice	75
BROWN, R. H., 'Empty-Field' Myopia and Visibility of Distant Objects at High Altitudes	376
BROWN, W. L., and OVERTON, R. K., Unrecognized Weakness and Compensatory Learning	126
BROWNSTEIN, A. J., and LAWSON, R., The Effect of Effort and Training-Test Similarity on Resistance to Extinction	123
BUCHWALD, A. M., The Generality of the Norms of Word-Associations	233
BURNHAM, R. W., Inter-Eye Differences in Color-Perception	386
BURT, H. E., and PRESSEY, S. L., Henry Herbert Goddard: 1866-1957	656
CAMPBELL, D. T., HUNT, W. A., and LEWIS, N. A., The Effects of Assimilation and Contrast in Judgments of Clinical Materials	347
CAPALDI, E. J., The Effect of Different Amounts of Alternating Partial Reinforcement on Resistance to Extinction	451
CHAPANIS, A., A Rate of Making Complex Decisions	650
CHO, J. B., and DAVIS, R. T., Preferences of Monkeys for Objects other than Food	87
CLARK, R. E., and HASLERUD, G. M., On the Redintegrative Perception of Words	97

CLAUSEN, J., and VANDERBILT, C., Visual Beats Caused by Simultaneous Visual and Photic Stimulation	577
COFER, C. N., and FORD, T. J., Verbal Context and Free Association-Time	606
COHEN, W., Spatial and Textural Characteristics of the <i>Ganzfeld</i>	403
COLE, D. L., Thirty-Seventh Annual Meeting of the Western Psychological Association	654
CONKLIN, J. E., Autonomous Changes in the Memory Trace	132
CONKLIN, J. E., The Influence of Figural Inspection on the Autokinetic Illusion	395
CROCKETT, W. H., and WEISSMAN, S. L., The Intersensory Transfer of Verbal Material	283
DALLENBACH, K. M., An Acknowledgment	471
DALLENBACH, K. M., Errata	471, 655
DALLENBACH, K. M., BITTERMAN, M. E., and NEWMAN, E. B., Editorial Announcements	135
DAVIS, R. T., and CHO, J. B., Preferences of Monkeys for Objects other than Food	87
DAVITZ, J. R., MASON, D. J., MOWRER, O. H., and VIEK, P., Conditioning of Fear: A Function of the Delay of Reinforcement	69
DAY, W. F., A Descriptive Analysis of the Non-Randomness of Serial Threshold-Data	238
DEESE, J., and MARDER, V. J., The Pattern of Errors in Delayed Recall of Serial Learning after Interpolation	594
DINNERSTEIN, D., and WERTHEIMER, M., Some Determinants of Phenomenal Overlapping	21
DUDEK, F. J., and BAKER, K. E., Scaling Line-Lengths with a Modification of the Constant-Sum Method	81
DUDEK, F. J., and BAKER, K. E., On the Validity of the Point-Assignment Procedure in the Constant-Sum Method	268
DUKES, W. F., An Inexpensive Variable-Pattern Maze	306
DUNCAN, C. P., Visual and Kinesthetic Components of Reactive Inhibition ..	616
EDWARDS, R. V., and GOTTSCHNEIDER, R. M., The Prediction of Collision ...	110
FALK, J. L., O'KELLY, L. I., and RIDGLEY, M., A Lever-Pressing Device for Fluid Rewards	303
FISKE, D. W., Twenty-Ninth Annual Meeting of the Midwestern Psychological Association	469
FORD, T. J., and COFER, C. N., Verbal Context and Free Association-Time ..	606
FORREST, D. W., Auditory Familiarity as a Determinant of Visual Threshold ..	634
FRANKS, C. M., and GRANGER, G. W., Dark-Adaptation and Some Observed Correlations	462
FRICK, J. W., and GUILFORD, J. P., An Analysis of a Form of the Water Jar Test	427
GARDNER, M. L., MICHAEL, A. L., ROBERTS, H. C., and STAGNER, R., A Continuous Millisecond Control for the Gerbrands Tachistoscope	299
GARDNER, R. A., Probability-Learning with Two and Three Choices	174
GIBSON, J. J., Technical and Scientific Communication: A Reply to Calvert ..	129
GIBSON, J. J., JOHNSON, C. W., STEINSCHEIDER, A., The Relative Accuracy of Visual Perception of Motion during Fixation and Pursuit	64
GIBSON, J. J., and SMITH, O. W., Apparatus for the Study of Visual Translatory Motion	291
GLEITMAN, H., Proactive and Retroactive Assimilation in the Successive Comparison of Loudness	117
GOTTSCHNEIDER, R. M., and EDWARDS, R. V., The Prediction of Collision	110
GRANGER, G. W., and FRANKS, C. M., Dark-Adaptation and Some Observed Correlations	462
GRUBER, H. E., and BERNARDIN, A. C., An Auditory Autokinetic Effect	133

GUILFORD, J. P., and FRICK, J. W., An Analysis of a Form of the Water Jar Test	427
GULICK, W. L., and STAKE, R. E., The Effect of Time on Size-Constancy	276
GUTTMANN, H. E., Constant-Current Sine-Wave Circuits for Stimulation	453
GUTTMANN, H. E., A Voice- or Sound-Key	456
HASLERUD, G. M., and CLARK, R. E., On the Redintegrative Perception of Words	97
HASTORF, A. H., and KENNEDY, J. L., Emmert's Law and Size-Constancy	114
HEIMER, W., and ROCK, I., The Effect of Retinal and Environmental Factors upon the Perception of Form	493
HERON, W., Perception as a Function of Retinal Locus and Attention	38
HESSERT, R. B., BRAUN, H. W., and ROCKLYN, E. H., Calibrated Materials for Verbal Learning in Middle- and Old-Aged Subjects	628
HILGARD, E. R., Lewis Madison Terman: 1877-1956	472
HOHLE, R., and PARDUCCI, A., Restriction of Range in the Judgment of Single Stimuli	272
HUNT, W. A., CAMPBELL, D. T., and LEWIS, N. A., The Effects of Assimilation and Contrast in Judgments of Clinical Materials	347
HURVICH, M. S., SHIPLEY, T. E., and WISHNER, J., The Serial-Position Curve as a Function of Organization	258
JAHNKE, J. C., A Group-Method for the Study of Figural After-Effects	319
JENKINS, W. L., and KARR, A. C., Paradoxical Warmth: A Sufficient Condition for Its Arousal	640
JOHNSON, C. W., GIBSON, J. J., SMITH, O. W., and STEINSCHNEIDER, A., The Relative Accuracy of Visual Perception of Motion during Fixation and Pursuit	64
JONES, F. N., An Analysis of Individual Differences in Olfactory Thresholds	227
KARLIN, L., and BRENNAN, G., Memory for Visual Figures by the Method of Identical Stimuli	248
KARR, A. C., and JENKINS, W. L., Paradoxical Warmth: A Sufficient Condition for Its Arousal	640
KASWAN, J., Association of Nonsense-Figures as a Function of Fittingness and Intention to Learn	447
KENNEDY, J. L., and HASTORF, A. H., Emmert's Law and Size-Constancy	114
KLEMMER, E. T., Rhythmic Disturbances in a Simple Visual-Motor Task	56
KOLERS, P. A., Subliminal Stimulation in Problem-Solving	437
LANE, G., Twenty-Eighth Annual Meeting of the Eastern Psychological Association	468
LANGHORNE, M. C., Third Annual Meeting of the Southwestern Psychological Association	321
LAWSON, R., and BROWNSTEIN, A. J., The Effect of Effort and Training-Test Similarity on Resistance to Extinction	123
LEIBOWITZ, H., and ZEIGLER, H. P., Apparent Visual Size as a Function of Distance for Children and Adults	106
LEWIS, N. A., CAMPBELL, D. T., and HUNT, W. A., The Effects of Assimilation and Contrast in Judgments of Clinical Materials	347
LOUTIT, C. M., The Use of Foreign Languages by Psychologists, Chemists, and Physicists	314
MAAG, C. H., Characteristics of Mental Impairment in Hypoxia	243
MARDER, V. J., and DEESE, J., The Pattern of Errors in Delayed Recall of Serial Learning after Interpolation	594
MASON, D. J., DAVITZ, J. R., MOWRER, O. H., and VIEK, P., Conditioning of Fear: A Function of the Delay of Reinforcement	69
MCGILL, T. E., and VERNON, J. A., The Effect of Sensory Deprivation upon Rote Learning	637
MELVILLE, J. R., Word-Length as a Factor in Differential Recognition	316
MESSICK, S. J., and SOLLEY, C. M., Probability-Learning, the Statistical Structure of Concepts, and the Measurement of Meaning	161
MESSICK, S. J., and SOLLEY, C. M., Word-Association and Semantic Differentiation	586

MEYER, M. F., The Tartini Pitches Created by the Primary Number Seven . . .	203
MEYER, M. F., Subjective Tones: Tartini and Beat-Tone Pitches	646
MICHAEL, A. L., GARDNER, M. L., ROBERTS, H. C., and STAGNER, R., A Continuous Millisecond Control for the Gerbrands Tachistoscope	299
MILLER, G. A., Some Effects of Intermittent Silence	311
MITCHELL, J. M., An Hypothesis of Psychological Aging	459
MONCRIEFF, R. W., Olfactory Adaptation and Odor-Intensity	1
MORSE, W. H., and SKINNER, B. F., A Second Type of Superstition in the Pigeon	308
MOWRER, O. H., DAVITZ, J. R., MASON, D. J., and VIEK, P., Conditioning of Fear: A Function of the Delay of Reinforcement	69
MULHOLLAND, T., The Frequency Procedures for Analyzing the Data Obtained by the Method of Limits	280
NACHMAN, M., The Influence of Size and Shape on the Discrimination of Visual Intensity	211
NAGEL, E. H., BATTIG, W. F., BROGDEN, W. J., and VOSS, J. F., Transfer and Retention of Bidimensional Compensatory Tracking after Extended Practice	75
NEISSER, U., Response-Sequences and the Hypothesis of the Neural Quantum	512
NEWMAN, E. B., Comment on the Information in Complex Decisions	652
NEWMAN, E. B., BITTERMAN, M. E., and DALLENBACH, K. M., Editorial Announcements	135
O'KELLY, L. I., FALK, J. L., and RIDGLEY, M., A Lever-Pressing Device for Fluid Rewards	303
OSGOOD, C. E., and ANDERSON, L., Certain Relations among Experienced Contingencies, Associative Structure, and Contingencies in Encoded Messages	411
OVERTON, R. K., and BROWN, W. L., Unrecognized Weakness and Compensatory Learning	126
PAPEZ, J. W., and PAPEZ, P., Mycotic Nature of Brain Damage in Mental Deficiency	333
PARDUCCI, A., Alternative Measures for the Discrimination of Shift in Reinforcement Ratio	194
PARDUCCI, A., and HOHLE, R., Restriction of Range in the Judgment of Single Stimuli	272
PASTORE, N., The Influence of Background on the Perception of Convexity and Concavity	131
POFFENBERGER, A. T., Harry Levi Hollingworth: 1880-1956	136
PORTER, L. W., The Effect of "Right" in a Modified Thorndikian Situation	219
PORTER, L. W., Effect of Shock-Cessation as an Incidental Reward in Verbal Learning	421
POSTMAN, L., Karen's Study of Incidental Learning as a Function of Meaningfulness	465
POSTMAN, L., and ADAMS, P. A., On Recent Studies of the Law of Effect in Incidental Learning	642
POULTON, E. C., Pursuit-Tracking with Partial Control of the Input	631
PRESSEY, S. L., and BURTT, H. E., Henry Herbert Goddard: 1866-1957	656
PRONKO, N. H., and WEHRENBURG, L., A Maze for Planaria	128
RAY, W. S., Verbal Compared with Manipulative Solution on an Apparatus-Problem	289
RIDGLEY, M., FALK, J. L., and O'KELLY, L. I., A Lever-Pressing Device for Fluid Rewards	303
ROBERTS, H. C., GARDNER, M. L., MICHAEL, A. L., and STAGNER, R., A Continuous Millisecond Control for the Gerbrands Tachistoscope	299
ROCK, I., The Role of Repetition in Associative Learning	186
ROCK, I., and HEIMER, W., The Effect of Retinal and Environmental Factors upon the Perception of Form	493
ROCKLYN, E. H., BRAUN, H. W., and HESSERT, R. B., Calibrated Materials for Verbal Learning with Middle- and Old-Aged Subjects	628
ROSEN, E., and RUSSELL, W. A., Frequency-Characteristics of Successive Word-Association	120

RUSSELL, W. A., and ROSEN, E., Frequency-Characteristics of Successive Word-Association	120
SALTZMAN, I. J., Incidental and Intentional Memory for Lifted-Weights	253
SARASON, B. R., and SARASON, I. G., Effects of Motivating Instructions and Reports of Failure on Verbal Learning	92
SARASON, I. G., The Effect of Associative Value and Differential Motivating Instructions on Serial Learning	620
SARASON, I. G., and SARASON, B. R., Effects of Motivating Instructions and Reports of Failure on Verbal Learning	92
SHAKLEE, A. B., Measurement of Cutaneous Electrical Thresholds in Animal Research	624
SHERLOCK, L., and SMITH, O. W., A New Explanation of the Velocity Transposition Phenomena	102
SHIPLEY, T. E., HURVICH, M. S., and WISHNER, J., The Serial-Position Curve as a Function of Organization	258
SKINNER, B. F., and MORSE, W. H., A Second Type of Superstition in the Pigeon	308
SMITH, O. W., and GIBSON, J. J., Apparatus for the Study of Visual Translatory Motion	291
SMITH, O. W., GIBSON, J. J., JOHNSON, C. W., and STEINSCHNEIDER, A., The Relative Accuracy of Visual Perception of Motion during Fixation and Pursuit	64
SMITH, O. W., and SHERLOCK, L., A New Explanation of the Velocity Transposition Phenomena	102
SMITH, O. W., and SMITH, P. C., Interaction of the Effects of Cues Involved in Judgments of Curvature	361
SOLLEY, C. M., and MESSICK, S. J., Probability-Learning, the Statistical Structure of Concepts, and the Measurement of Meaning	161
SOLLEY, C. M., and MESSICK, S. J., Word-Association and Semantic Differentiation	586
STAATS, A. W., Verbal and Instrumental Response-Hierarchies and Their Relationship to Problem-Solving	442
STAGNER, R., GARDNER, M. L., MICHAEL, A. L., and ROBERTS, H. C., A Continuous Millisecond Control for the Gerbrands Tachistoscope	299
STAKE, R. E., and GULICK, W. L., The Effect of Time on Size-Constancy	276
STANLEY, J. C., Index of Means vs. Mean of Indices	467
STEINSCHNEIDER, A., GIBSON, J. J., JOHNSON, C. W., and SMITH, O. W., The Relative Accuracy of Visual Perception of Motion during Fixation and Pursuit	64
STEVENS, J. C., and TULVING, E., Estimation of Loudness by a Group of Untrained Observers	600
STEWART, P. A., and THOMAS, G. J., The Effect on Visual Perception of Stimulating the Brain with Polarizing Currents	528
TALLAND, G. A., Criteria in Conceptual Transposition	263
THOMAS, G. J., and STEWART, P. A., The Effect on Visual Perception of Stimulating the Brain with Polarizing Currents	528
THURLOW, W., An Auditory Figure-Ground Effect	653
TULVING, E., and STEVENS, J. C., Estimation of Loudness by a Group of Untrained Observers	600
VANDERBILT, C., and CLAUSEN, J., Visual Beats Caused by Simultaneous Visual and Photoc Stimulation	577
VAUGHAN, C. J., and BENDIG, A. W., Manifest Anxiety, Discrimination, and Transposition	286
VERNON, J. A., and MCGILL, T. E., The Effect of Sensory Deprivation upon Rote Learning	637
VIEK, P., DAVITZ, J. R., MASON, D. J., and MOWRER, O. H., Conditioning of Fear: A Function of the Delay of Reinforcement	69
VOSS, J. F., BATTIG, W. F., BROGDEN, W. J., and NAGEL, E. H., Transfer and Retention of Bidimensional Compensatory Tracking after Extended Practice	75

WATSON, W. A., Contrast, Assimilation, and the Effect of Central Tendency . .	560
WAYNER, M., An Inexpensive Respirator for Use with Small Animals	457
WEHRENBURG, L., and PRONKO, N. H., A Maze for Planaria	128
WEINSTEIN, S., The Perception of Depth in the Absence of Texture-Gradient	611
WEISSMAN, S. L., and CROCKETT, W. H., Intersensory Transfer of Verbal Material	283
WERTHEIMER, M., and DINNERSSTEIN, D., Some Determinants of Phenomenal Overlapping	21
WISNER, J., SHIPLEY, T. E., JR., and HURVICH, M. S., The Serial-Position Curve as a Function of Organization	258
WODINSKY, J., and BITTERMAN, M. E., Discrimination-Reversal in the Fish	569
WOLFF, W., Fourth Interamerican Congress of Psychology	470
YOUNG, P. T., Continuous Recording of the Fluid Intake of Small Animals . .	295
ZEIGLER, H. P., and LEIBOWITZ, H., Apparent Visual Size as a Function of Distance for Children and Adults	106

BOOK REVIEWS

(The reviewer's name appears in parentheses after the title of the work.)

AMBROSE, G., and NEWBOLD, G., A Handbook of Medical Hypnosis: An Introduction for Practitioners and Students (F. L. Marcuse)	661
ANSBACHER, H. L., and ANSBACHER, R. R., Editors, The Individual Psy- chology of Alfred Adler (I. E. Alexander)	151
BLAIR, G. W. S., Measurement of Mind and Matter (K. E. Baker)	329
BLUM, M. L., Industrial Psychology and Its Social Foundations (R. S. Harper)	159
BOND, G. L., and TINKER, M. A., Reading Difficulties—Their Diagnosis and Correction (K. Ware)	332
BROWN, C. W., and GHISELLI, E. E., Scientific Method in Psychology (R. E. Dreher)	155
BRENNER, J. S., SMITH, M. B., and WHITE, R. W., Opinions and Personality (R. D. Walk)	665
BRUNSWIK, E., Perception and the Representative Design of Psychological Experiments (J. E. Hochberg)	480
BURTON, A., and HARRIS, R. E., Editors, Clinical Studies of Personality (J. Wishner)	145
BUXTON, C. E., College Teaching: A Psychologist's View (L. M. Smith) . . .	677
CARROLL, J. B., Editor, Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf (G. A. Miller)	152
CHERRY, C., On Human Communication: A Review, a Survey, and a Criticism (G. A. Miller)	322
CLARK, F. L., and DUNNE, A. C., Ageing in Industry (H. Meltzer)	667
CRUICKSHANK, W., Editor, Psychology of Exceptional Children (E. A. Doll)	153
DARLEY, J. G., and HAGENAH, T., Vocational Interest Measurement: Theory and Practice (T. A. Ryan)	325
DAVID, H. P., and VON BRACKEN, H., Editors, Perspectives in Personality (I. E. Alexander)	664
DORCUS, R., Editor, Hypnosis and Its Therapeutic Applications (F. L. Marcuse)	661
DUNNE, A. C., and CLARK, F. L., Ageing in Industry (H. Meltzer)	667
FAIRFIELD, L., Epilepsy: Grand Mal, Petit Mal Convulsions (F. A. Gibbs) . .	670
FARNSWORTH, P. R., Editor, and MCNEMAR, O., Associate Editor, Annual Review of Psychology (F. N. Jones and M. H. Jones)	323
FESTINGER, L., RIECKEN, H. W., and SCHLACHTER, S., When Prophecy Fails (A. Pepitone)	671
GHISELLI, E. E., and BROWN, C. W., Scientific Method in Psychology (R. E. Dreher)	155
GLICK, P. C., American Families: A Demographic Analysis of Census Data on American Families at Mid-Century (M. B. Drury)	676

HAGENAH, T., and DARLEY, J. G., Vocational Interest Measurement: Theory and Practice (T. A. Ryan)	325
HARRIS, R. E., and BURTON, A., Editors, Clinical Studies of Personality (J. Wishner)	145
HENRY, W. E., The Analysis of Fantasy: The Thematic Apperception Technique in the Study of Personality (T. E. Sarbin)	158
IKIN, A. G., New Concepts of Healing: Medical, Psychological, and Religious (D. Lebo)	487
JONES, M. R., Nebraska Symposium on Motivation: 1955 (S. Feshbach)	147
KAPLAN, O. J., Editor, Mental Disorders in Later Life (C. Landis)	156
KATZ, E., and LAZARSFELD, P. F., Personal Influence: The Part Played by People in the Flow of Mass Communications (A. Pepitone)	157
KLEIN, J., The Study of Groups (H. Winthrop)	668
LEVINE, E. S., Youth in a Soundless World: A Search for Personality (H. S. Lane)	663
LOFQUIST, L. H., Vocational Counseling with the Physically Handicapped (S. G. DiMichael)	659
MCNEMAR, O., Associate Editor, and FARNSWORTH, P. R. Editor, Annual Review of Psychology (F. N. Jones and M. H. Jones)	323
MICHAELIS, A. R., Research Films in Biology, Anthropology, Psychology, and Medicine (C. R. Carpenter)	156
MINER, J. B., Intelligence in the United States: A Survey—with Conclusions for Manpower Utilization in Education and Employment (K. Bradway) ..	668
MURPHY, L. B., <i>et al.</i> , Personality in Young Children (P. Mussen)	327
NEWBOLD, G., and AMBROSE, G., A Handbook of Medical Hypnosis: An Introduction for Practitioners and Students (F. L. Marcuse)	661
RIECKEN, H. W., FESTINGER, L., and SCHACHTER, S., When Prophecy Fails (A. Pepitone)	671
ROBACK, A. A., Freudiana: Including Unpublished Letters from Freud, Havlock Ellis, Pavlov, Bernard Shaw, Romain Rolland, <i>et alii</i> (S. Rosenzweig) ..	669
ROE, A., The Psychology of Occupations (R. S. Harper)	663
SCHACHTER, S., FESTINGER, L., and RIECKEN, H. W., When Prophecy Fails (A. Pepitone)	671
SCHILDER, P., The Nature of Hypnosis (F. L. Marcuse)	661
SHOLL, D. A., The Organization of the Cerebral Cortex (J. W. Papez)	673
SMITH, M. B., BRUNER, J. S., and WHITE, R. W., Opinions and Personality (R. W. Walk)	665
SNYDER, L. L., The Meaning of Nationalism (K. F. Muenzinger)	154
SOROKIN, P. A., Fads and Foibles in Modern Sociology and Related Sciences (G. B. Vetter)	674
SPENCE, K. W., Behavior Theory and Conditioning (M. E. Bitterman)	141
TINKER, M. A., and BOND, G. L., Reading Difficulties—Their Diagnosis and Correction (K. Ware)	332
TYLER, L. E., Editor, The Psychology of Human Differences (J. Brözek) ...	149
UNDERWOOD, B. J., Psychological Research (W. F. Grether)	658
VAN PELT, S. J., Hypnotic Suggestion: Its Role in Psychoneurotic and Psychosomatic Disorders (F. L. Marcuse)	661
VARIOUS AUTHORS, A Twentieth Century Bestiary (R. Held)	331
VARIOUS AUTHORS, <i>L'Analyse Factorielle et ses Applications</i> (W. B. Michael) ..	330
VERNON, M. D., Backwardness in Reading: A Study of Its Nature and Origin (M. A. Tinker)	675
VON BRACKEN, H., and DAVID, H. P., Editors, Perspectives in Personality (I. E. Alexander)	664
WHITE, R. W., BRUNER, J. S., and SMITH, M. B., Opinions and Personality (R. D. Walk)	665
WILLIAMS, R. J., Biochemical Individuality, The Basis for the Genetotrophic Concept (C. M. McCay)	672
YATABE, T., <i>Shiko Shinrigaku</i> (The Psychology of Thinking III—Thinking in Animals (B. Yagi)	485

THE AMERICAN JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXX

MARCH, 1957

No. 1

OLFACTORY ADAPTATION AND ODOR-INTENSITY

By R. W. MONCRIEFF, Chichester, England

The concept of odor-intensity is one that is familiar; we easily recognize the smells of *n*-butanol and amyl acetate as intense and those of methanol and coumarin as relatively weak. Despite the general familiarity with the concept, there appears, however, to be no generally accepted definition of odor-intensity; it is with the provision of such a definition and with the establishment of a relation between the so-defined intensity of odor and the concentration of the odorant in solution, that this paper is concerned.

The most important of the older work that has been published on odor-intensity would appear to be that of Zwaardemaker and of Fair and Wells.¹ Zwaardemaker devised an olfactometer in which air was passed through a tube over an odorous surface and in which the area of this surface could be varied; when the odorized area to which the air was exposed was one, two, or three units then the issuing air was said to have one, two, or three *olfacties* or units of stimulus-odor.

Fair and Wells devised an osmoscope which has been widely used to measure the odors of samples of impure water; the principle on which it works is that the observer measures the number of times that a sample of the air standing over the water must be diluted with odorless air to reach its odor-threshold. Many other olfactometers that depend on the same principle of progressive dilution until the odor-threshold is reached have been made. One of the best known of these was the Allison and Katz odorometer which in an improved form was used by the U. S. Bureau of Mines in a study of the use of odorants such as ethyl mercaptan and pyridine as warning agents in case of leakage of natural fuel gases which did

* Received for publication March 29, 1956. This study was carried out with funds provided by Airkem, Inc., New York City, New York.

¹ Hendrik Zwaardemaker, *Die Physiologie des Menschen*, 1891, 1-324; Ein verbesserter Riechmesser, *Arch. Laryngol. Rhinol.*, 3, 1895, 368-371; G. M. Fair and W. F. Wells, Air dilution method of odor determination in water analysis, *J. Amer. Water Works Assn.*, 1933, 1670-1677; Odor observer and odor meter, U. S. Patent 2136844 (1939).

Dated

Accs. No.

not have an intrinsic powerful odor.² While such instruments are undoubtedly useful, they cannot, however, be said to offer a measure of the intensity of an odor.

An attempt was made by Beck, Kruger, and Calabresi to compare the odor-intensities of different substances with the odor-intensity of an arbitrarily selected standard odorant which was heptaldehyde in a series of dilutions in benzyl benzoate.³ This is something like the problem of determining when a red light is equal in intensity to a white light; it requires the observer to do something that he does not customarily do and that he may not have the mental equipment to do; if he is to pretend to be a monochrome photographic film it is going to make a world of difference whether he decides to be a 'pan' film or merely 'ortho.' When the comparisons of smell-intensities were made by Beck, Kruger, and Calabresi, they found that reproducibility in any one individual observer was good, but that there were such great differences between the results of three different and highly-trained observers that they did not even attempt to make any assessment of group response. Probably the three observers were comparing not the same, but different, properties.

FACTORS ASSOCIATED WITH INTENSITY OF ODORS

Three properties seem to be quite closely associated with odor-intensity. They are: (1) a low threshold concentration; (2) a high vapor pressure; and (3) masking ability.

(1) *A low threshold concentration.* Substances such as the mercaptans and pyridines which have intense odors have low threshold values. There are, though, other substances such as acetone which have intense odors and which do not have very low threshold concentrations, and contrarily the musks have extremely low threshold values, but their smells, although very persistent and pervasive, cannot be described as intense. Musks are commonly added to perfumes to increase their persistence and to 'round off' the odors of the constituents, but the smell of musk is not as a rule noticeable in the finished perfume; it is not intense enough. A number of workers have made measurements of the threshold concentrations of odorants, and the values for a fairly large number of substances are known, although there are often considerable discrepancies between the results of different workers.

(2) *A high vapor pressure.* Some of the intense odorants such as ether, carbon disulphide, acetone, and chloroform have very high vapor pressures, from 150 to 450 mm. mercury at 20°C; on the other hand, geranyl acetate and citral which both have intense smells have vapor pressures about 10,000 times lower according to values given by Crocker.⁴ There are, too, odorous substances such as methanol which have a high vapor pressure but only a weak smell. It is relevant that, following an unsatisfactory attempt by Kenneth to relate the odor-intensities of san-

² V. C. Allison and S. H. Katz, An investigation of stenches and odors for industrial purposes, *J. industr. eng. Chem.*, 11, 1919, 336-338.

³ L. H. Beck, Lawrence Kruger, and P. Calabresi, Observations on olfactory intensity, *Ann. N. Y. Acad. Sci.*, 58, 1954 (No. 2), 225-238.

⁴ E. C. Crocker, Enter—the engineer: Scientific principles and how they can be applied to the perfumer's art, *Amer. Perfumer Essential Oil Rev.*, 31, 1935 (October) 83-86.

dalwood, cedar, origanum, and terebene to their specific gravities,⁵ Durrans suggested that intensity of smell could better be related to volatility,⁶ which Kenneth seems to have accepted.⁷ The relation between the vapor pressures and odor-intensities of the aliphatic alcohols was studied by Kruger, Feldzamen and Miles.⁸ They found that high volatility contributes to the intensity of odor, but apparently has only a secondary influence. There is, however, no doubt that high volatility is often associated with intensity of odor, but there is nothing to support a hypothesis that odor-intensity is entirely dependent on high volatility.

(3) *Masking ability.* Some substances which have intense smells will mask, or obliterate other smells; amyl butyrate, for example, is sometimes used to cover unpleasant odors in industrial products. Most masking agents do have intense smells and usually, too, they have low threshold concentrations; their volatility is more variable. 'Masking' is a term that has been used by different people with different shades of meaning. The usually accepted definition, that given by McCord and Witheridge, is: "odor-masking is the process of obliterating one odor by superimposing another to create a more compelling sensation, preferably pleasant."⁹

Lowness of threshold concentration alone is insufficient to ensure intensity of odor although it is often associated with it; volatility is helpful but not more closely tied to intensity, and masking ability seems to be closely associated with intensity.

During a series of experiments on olfactory adaptation extending over a period of about six months, the effect of first inspiring an odorant on the threshold concentration of that odorant was repeatedly investigated (adaptation in a homogeneous pair). It was found that the adaptation resulted in a considerable enhancement of the normal threshold concentration. The figures obtained from the experiments were treated in various ways; for one purpose the enhancement, *i.e.* number of times that the threshold concentration of a particular odorant was raised by a previous inspiration of the undiluted odorant, was calculated. Inspection of the results at once suggested a close relationship between the degree of this enhancement and the intensity of the odorant, and experiments were then made specifically to investigate this relationship.

METHOD

(1) Observations were made by smelling 25 ml. of the odorant in a 200 ml. wide-neck bottle. The observer (*O*) sat on a stool near a slightly open window with eyes closed; the neck of a bottle containing undiluted odorant was just

⁵ J. H. Kenneth, Smell and specific gravity, *Nature* (London), 111, 1923, 151.

⁶ T. H. Durrans, Factors of odour strength, *ibid.*, 111, 1923, 359.

⁷ Kenneth, Odours and the sense of smell, *ibid.*, 117, 1926, 591 f.

⁸ Lawrence Kruger, A. N. Feldzamen, and W. S. Miles, Comparative olfactory intensities of the aliphatic alcohols, this JOURNAL, 68, 1955, 386-395.

⁹ C. P. McCord and W. N. Witheridge, *Odors: Physiology and Control*, 1949, 186.

touched in the angle between his upper lip and his nose and *O* inspired once, deliberately but not especially deeply, while the bottle was held there. The bottle was removed and *O* breathed out; another bottle containing a dilute solution of the same odorant was touched under his nose and he inspired again. This time he said 'yes' or 'no' according as he could, or could not, smell the odorant. Thus, in his first breath *O* was adapted with the undiluted odorant and in his second breath he inspired the test-solution.

On another series of tests the first bottle offered contained not the odorant but the diluent, and the second bottle, as before, a dilute solution of the odorant. In these two ways responses were obtained for: (a) test-solution of odorant in diluent after smelling undiluted odorant; and (b) test-solution of odorant in diluent after smelling diluent alone. Even although the diluent was often water which *O* could not smell, the same routine was maintained; at the least, it gave *O* a reference or control with which to compare the smell of the test-solution.

Definite responses were encouraged but *O* was allowed to say 'doubtful' when undecided. Requests for repeat tests were allowed only when something had gone wrong with the test, *e.g.* the bottle had been touched to the wrong part of the nose or out of phase with *O*'s breathing; they were not allowed for indecision, simply to let *O* make up his mind. In some respects the experimental method was similar to that described by Cheesman and Mayne.¹⁰ Four *O*s made tests. They were *RWM*, a 52-yr.-old man; *RMM*, *SL*, and *ST*, women, 24, 17, and 16 yr. old respectively. *RMM* is the daughter of *RWM*. All of the *O*s were known from previous work to have a reasonably uniform and normal sense of smell. *RWM* and *SL* took part in all of the tests, *RMM* took part in the acetone tests and *ST* in the others. After every four tests, *O* was rested for several minutes, usually *O* and *E* exchanging roles. In those tests involving the use of methanol and ethanol it was found necessary to rest *O* for a few minutes after each and every test. This was not necessary as indicated by the reproducibility of results with any of the other odorants used. Scoring was done on the basis of 1 point for a 'yes,' 0.5 points for a 'doubtful,' and nil for a 'no' response. The percentage positive responses for each test concentration offered was calculated, and the lowest concentration (which we designated PR_{50}) at which the group scored 50%, *i.e.* gave at least 50% positive responses, was determined. For those odorants that were soluble in it, distilled water was the chosen diluent; for the others propylene glycol (1:2 propane-diol), which had only a negligible smell, was used. Dilute solutions were freshly prepared twice each day from stronger stock solutions. Interspersed among the test-dilutions offered to *O* there were always tests of the diluent alone, so that *O* never knew whether the test-solution he was smelling contained or did not contain any odorant. Usually about six different concentrations would be offered in random order, after a few pilot runs had first been made, to find the approximate threshold, and each concentration was offered five times, as was also the diluent. The precautions taken had the effect of making *O* treat each smell on its own merits and give an objective response. Some 20 odorant materials were tested as described.

When the PR_{50} concentrations for (a) odorant smelled after undiluted odorant (C_1), and (b) odorant smelled after diluent (C_2) had been obtained, the value

¹⁰ G. B. Cheesman and S. Mayne, The influence of adaptation on absolute threshold measurements for olfactory stimuli, *Quart. J. exp. Psychol.*, 5, 1953, 22-30.

of the ratio C_1/C_2 was calculated and this ratio was taken to be the intensity of the odorant. It is, in effect, a measure of the enhancement of the threshold concentration which is caused by adaptation with the odorant itself. For low intensities, the expression C_1/C_2 approaches unity and the odor-intensity of any odorant with which previous smelling does not raise its own threshold value is unity. In practice only those odorants with the faintest of smells would have an intensity of unity; all odorants with easily noticeable smells cause some adaptation so that their intensity values will be higher than unity.

(2) The experimental method already described was extended to measure the intensities of solutions of different concentrations of (a) acetone; (b) *isopropanol*; and (c) ethyl mercaptan. In this series of tests a solution of odorant, e.g. 10% acetone in water, was considered as an 'undiluted odorant,' and adaptation was made with it. This was followed by smelling a test-solution consisting of a dilution of the 10% acetone; and similarly for other concentrations.

(3) When the results given by Method (2) had been examined, it could be seen that it might be easier to distinguish between different concentrations of an odorant in one range of concentrations than in another. Accordingly a series of concentrations of acetone, and later of *isopropanol*, and ethyl mercaptan were made up and 25 ml. of each introduced into 200 ml. wide-neck stoppered bottles which were coded. Then *O* was asked to place the bottles in order of smell-intensity. Three *Os*, *RWM*, *SL* and *ST*, took part in the acetone and some of the *isopropanol* tests and two, *RWM* and *SL*, in the ethyl mercaptan tests and in the other part of the *isopropanol* tests. The sorting of the bottles was done several times and for each run they were so re-coded that recollection of the results of the first run would not influence the second run and so on. A device which we used latterly and which saved the labor of re-coding was to stick small labels, of the kind used for microscope slides, pencilled with a code letter, on the bottom of the bottles; these could be read only by holding the bottles aloft and there was no chance of seeing them by accident.

RESULTS

(1) *Intensity of pure (undiluted) odorants: (a) Acetone.* The primary results are given in full for the first odorant used which was acetone, because they illustrate the spread of results within one individual's responses as well as the variations from one individual to another, and also because they will exemplify clearly the way in which the results were treated to obtain values of odor-intensity.

In the first series of tests the first bottle that was held to *O*'s nose contained undiluted acetone and the second bottle, for his next inspiration, contained either a solution of acetone in water or else water itself, as a reliability check. The results are given in Table I, a 'yes' being indicated as '+', a 'no' as '—', and a 'doubtful' response as '?'. It can be seen from Table I that the PR_{50} concentration of acetone for the group after adaptation with undiluted acetone is 5%. Next, test-solutions of acetone were

offered after 'adaptation' with the pure diluent, water, and the results obtained are shown in Table II, from which it can be seen that the PR_{50} concentration of acetone for the group after 'adaptation' with water is 0.03%. The ratio of these two threshold concentrations PR_{50} (after acetone)/ PR_{50} (after water) or $C_1/C_2 = 5/0.03 = 167$, and this is taken to be a measure of the intensity of the smell of acetone.

TABLE I

OLFACTORY RESPONSES TO SOLUTIONS OF ACETONE AFTER SMELLING UNDILUTED ACETONE WITH THE PREVIOUS INSPIRATION

Concentration of test-solution (% volume)	O			Positive responses of group (%)
	RWM	SL	RMM	
8	++++ ?	++++++	++++++	90
7	++-++	++++++	++++++	87
6	+---+	++++++	++++++	87
5	-++--	++++++	+++ ? -	70
4	-----+	---+--	---+-+	27
0 (water)	-----	-----	-----	0

TABLE II

OLFACTORY RESPONSES TO SOLUTIONS OF ACETONE AFTER 'SMELLING' THE DILUENT, WATER, WITH THE PREVIOUS INSPIRATION

Concentration of test-solution (% volume)	O			Positive responses of group (%)
	RWM	SL	RMM	
0.08	++++++	++++++	++++++	100
0.06	++ ? ++	++++++ ?	+---++	87
0.04	-+ ? + ?	? -++ ?	---+-+	53
0.03	+++--	? -++-	? -++-	53
0.02	---+-	+---+-	---+-	27
0.00 (water)	- ? ---	-----	-----	3

(b) Other odorants. When trials were similarly carried out with other odorant materials the results obtained were as shown in Table III in which acetone is also included for convenience of comparison.

(2) *Variation with concentration.* Acetone was the first odorant used and a series of concentrations from 0.02 to 100% was prepared. The PR_{50} threshold concentration, (C_1), after adaptation with each of these prepared concentrations was determined; the results are shown in Table IV. This table also shows the PR_{50} concentration (C_2) after 'adaptation' with water, and lastly the intensity (C_1/C_2) of the odor of each concentration.

Similarly, the results for different concentrations of isopropanol are shown in Table V and those for ethyl mercaptan in Table VI.

(3) *Arrangement of dilutions according to odor-intensity.* A series of dilutions of isopropanol in distilled water was prepared and coded. It

TABLE III
EVALUATION OF ODOR-INTENSITY OF A NUMBER OF ODORANTS

Odorant	Type of odor	Diluent	PR ₅₀ after undiluted odorant C ₁	PR ₅₀ after diluent C ₂	Intensity C ₁ /C ₂
Acetone	characteristic	water	5.0	0.03	167
Allyl caproate	pineapple (irritant concen- trated)	propylene glycol	0.1	0.00002	5000
Amyl acetate	pear drops	propylene glycol	0.3	0.002	150
Benzaldehyde	almonds, nutty	propylene glycol	0.5	0.05	10
Benzylamine	fishy, bitter, slightly spicy	water	0.5	0.3	1.7
n-Butanol	fusel-oil, burnt, persistent	water	1.0	0.005	200
iso-Butanol	bitter, fusel-oil, buttery note	water	0.3	0.02	15
sec-Butanol	fusel-oil, more astringent than n-butanol	water	0.2	0.005	40
tert-Butanol	reminiscent of mixture of iso-propanol and acetone; fusel-oil, slightly spicy and camphoraceous	water	2.0	0.1	20
Butyl acetate	pear drops, slightly sharper than amyl acetate	propylene glycol	0.5	0.03	17
n-Butyric acid	sour, bitter (irritant concen- trated)	water	0.1	0.001	100
Cellosolve	fishy, bitter	water	0.6	0.02	30
Clove oil	cloves	propylene glycol	0.5	0.003	170
Diacetone alcohol	nauseous, floral, fruity	water	1.0	0.01	100
Ethanol	spirituous	water	3.5	0.3	12
Ethyl mercaptan (0.5%)	foul, alliaceous	water	0.0028 (one observer only)	0.0000002	14000
α-Ionone	violets	propylene glycol	1.5	0.05	30
Methanol	wood spirit, slightly reminis- cent of shell fish	water	2.0	0.2	10
Nitrobenzene	almonds, organic	propylene glycol	0.5	0.05	10
n-Propanol	sweeter than iso-propanol no thyme suggestion	water	0.7	0.05	14
Isopropanol	spirituous, earthy reminis- cent of thyme	water	2.0	0.07	28
Pyridine	rank, gassy, tobacco-like	water	0.07	0.00005	1400

was thought likely that by taking sufficient time over the arrangement these dilutions could be arranged in correct order. One O (SL) essayed such an arrangement, but she required 3 hr. work to be satisfied that the

bottles were arranged properly. Her arrangement was as follows, in order of decreasing smell-intensity, all figures showing percentage volume of isopropanol: 100, (80, 90), 70, 60, (40, 45, 50), 35, 30, 25, 20, 15, 10,

TABLE IV

VARIATION OF INTENSITY OF ODOR WITH CONCENTRATION OF ACETONE SOLUTIONS IN WATER

Concentration of adapting solution (% volume)	PR ₅₀ after adaptation with acetone C ₁	PR ₅₀ after 'adaptation' with water C ₂	Intensity of acetone solution used for adaptation C ₁ /C ₂
100	5	0.03	167
80	6	0.03	200
75	5	0.03	167
62	5	0.03	167
50	5	0.03	167
40	4	0.03	133
25	3	0.03	100
10	1.5	0.03	50
5	1.0	0.03	33
3	0.7	0.03	23
1.5	0.5	0.03	17
1.0	0.4	0.03	13
0.5	0.3	0.03	10
0.2	0.14	0.03	5
0.1	0.10	0.03	3
0.05	0.07	0.03	2
0.02	0.03	0.03	1

TABLE V

VARIATION OF INTENSITY OF ODOR WITH CONCENTRATION OF ISOPROPANOL SOLUTIONS IN WATER

Concentration of adapting solution (% volume)	PR ₅₀ after adaptation with isopropanol C ₁	PR ₅₀ after 'adaptation' with water C ₂	Intensity of isopropanol solution used for adaptation C ₁ /C ₂
100	2.0	0.07	29
85	1.7	0.07	24
75	2.0	0.07	29
60	2.3	0.07	33
50	2.3	0.07	33
30	2.3	0.07	33
25	2.0	0.07	29
20	1.8	0.07	26
15	1.7	0.07	24
10	1.2	0.07	17
6	0.8	0.07	11
3	0.5	0.07	7
1	0.3	0.07	4
0.6	0.2	0.07	3
0.3	0.15	0.07	2
0.1	0.1	0.07	1.4

5, 1, 0.6, 0.3, 0.1, 0 (water). Apart from being unable to separate 80 and 90% which were grouped together, and 40, 45 and 50% also grouped together, the arrangement was perfect.

The point at issue was, however, not so much the possibility of sorting out such dilutions by smell, as the relative ease of sorting in different ranges of concentration. The test had to be made so difficult that mistakes would be made, and then the incidence of the mistakes in different con-

TABLE VI
VARIATION OF INTENSITY OF ODOR WITH CONCENTRATION OF
ETHYL MERCAPTAN SOLUTIONS IN WATER

Concentration of adapting solution (% volume)	PR ₅₀ after adaptation with ethyl mercaptan C ₁	PR ₅₀ after 'adaptation' with water C ₂	Intensity of ethyl mercaptan solution used for adaptation C ₁ /C ₂
0.5	0.0028 (one observer only)	0.0000002	14000
0.3	0.0020	0.0000002	10000
0.2	0.0015	0.0000002	7500
0.1	0.0010	0.0000002	5000
0.05	0.0005	0.0000002	2500
0.005	0.00015	0.0000002	750
0.0005	0.00005	0.0000002	250
0.00005	0.00001	0.0000002	50
0.000005	0.000002	0.0000002	10
0.0000005	0.0000005	0.0000002	2.5
0.0000002	0.0000002	0.0000002	1

TABLE VII
ORGANOLEPTIC APPRAISAL OF SMELL INTENSITY OF DILUTIONS OF ISOPROPANOL
Positions assigned to coded solutions

Concentration isopropanol in water (% volume)	RWM				SL				ST				Group mean	'Correct' position	SD from 'correct' position	Coefficient of variation (% from 'correct' position)
	1st run	2nd run	3rd run	4th run	1st run	2nd run	3rd run	4th run	1st run	2nd run	3rd run	4th run				
100	1	5	9	1	1	1	1	1	1	1	7	2	2.58	1	3.12	312
90	3	3	7	2	7	3	7	7	3	2	5	8	4.75	2	3.52	176
80	2	2	8	10	6	5	9	6	2	7	3	4	5.33	3	3.35	112
70	7	10	6	5	2	9	2	4	5	6	1	1	4.83	4	2.90	73
60	11	6	10	6	10	7	6	9	7	9	4	5	7.50	5	3.29	66
50	10	8	11	4	4	4	11	5	11	4	8	3	6.92	6	3.23	54
45	8	9	2	8	3	10	5	3	6	10	6	6	6.33	7	2.71	39
40	9	1	3	7	12	2	10	10	4	3	2	9	6.0	8	4.22	53
35	12	4	1	9	9	6	3	8	10	5	10	7	7.0	9	3.72	41
30	5	7	5	12	11	11	4	2	8	8	9	10	7.7	10	3.81	38
25	4	13	12	13	14	13	15	11	9	13	11	13	11.8	11	2.87	26
20	13	12	13	11	8	8	12	14	13	11	13	12	11.7	12	1.87	16
15	6	14	4	3	5	15	8	13	12	12	14	11	9.7	13	5.25	40
10	14	11	14	15	13	12	13	12	14	14	12	14	13.2	14	1.41	10
5	15	15	15	14	15	14	14	15	15	15	15	15	14.8	15	0.50	3
1	—	—	18	18	—	—	19	16	—	—	16	16	17.2	16	1.68	10
0.6	—	—	16	16	—	—	16	19	—	—	17	17	16.8	17	1.08	6
0.3	—	—	17	19	—	—	18	18	—	—	19	18	18.2	18	0.71	4
0.1	—	—	19	17	—	—	17	17	—	—	18	19	17.8	19	1.47	8
0.0	20	20	20	20	20	20	20	20	20	20	20	20	20	20	0	0

centration ranges could be examined; accordingly each *O* was limited in subsequent experiments to a maximum of 30 min. to make each arrangement.

When this was done the results shown in Table VII were obtained; the placing given for each of the concentrations is shown. In the first run by

RWM, 100% was placed first in intensity, 80% second, 90% third, 25% fourth, 30% fifth, and so on. The mean for the group of three *O*s was calculated, and also the standard deviation (*SD*) from the 'correct' position for each concentration, and the coefficient of variation from the

TABLE VIII
ORGANOLEPTIC APPRAISAL OF SMELL INTENSITY OF DILUTIONS OF ISOPROPANOL

Concentration isopropanol in water (% volume)	Positions assigned to coded solutions										Group mean	'Correct' position	SD from 'correct' position	Coefficient of variation (%) from 'correct' position
	RWM					SL								
	1st run	2nd run	3rd run	4th run	5th run	1st run	2nd run	3rd run	4th run	5th run				
100	1	1	1	1	1	1	1	1	1	1	1.0	1	0	0
90	10	3	4	5	2	8	3	7	4	8	5.4	2	4.25	212
80	6	5	3	6	4	4	10	10	3	4	5.5	3	3.51	117
70	9	9	7	8	11	7	5	6	9	7	7.8	4	4.15	106
60	12	4	2	4	3	9	12	4	8	9	6.7	5	3.94	79
50	11	7	5	7	6	6	11	9	2	2	6.6	6	3.07	51
45	2	6	8	2	8	3	2	11	7	3	5.2	7	3.55	51
40	8	14	6	3	5	5	9	12	11	10	8.3	8	3.36	42
35	4	13	11	10	10	2	6	8	6	11	8.1	9	3.45	38
30	7	2	9	11	12	12	7	3	10	5	7.8	10	4.07	41
25	3	8	10	9	7	10	4	2	5	6	6.4	11	5.35	49
20	5	11	13	12	13	11	8	5	13	13	10.4	12	3.46	29
15	13	10	12	13	9	13	13	13	12	12	12.0	13	1.67	13
10	14	12	14	14	14	14	15	14	14	15	14.0	14	0.77	6
5	15	15	15	15	15	15	14	15	15	14	14.8	15	0.45	3
1	16	16	16	18	16	16	16	16	16	16	16.2	16	0.45	3
0.6	17	17	17	16	17	17	18	18	17	17	17.1	17	0.55	3
0.3	19	20	18	17	18	18	17	17	18	18	18.0	18	0.92	5
0.1	20	19	19	20	19	19	19	19	19	19	19.2	19	0.45	2
0.0	18	18	20	19	20	20	20	20	20	20	19.5	20	0.95	5

TABLE IX
ORGANOLEPTIC APPRAISAL OF SMELL INTENSITY OF DILUTIONS OF ACETONE

Concentration acetone in water (% volume)	Positions assigned to coded solutions									Group mean	'Correct' position	SD from 'correct' position	Coefficient of variation (%) from 'correct' position
	RWM			SL			ST						
	1st run	2nd run	3rd run	1st run	2nd run	3rd run	1st run	2nd run	3rd run				
100	2	1	2	3	1	4	2	3	2	2.22	1	1.52	152
90	4	2	1	1	5	1	4	1	6	2.78	2	2.03	102
80	3	3	4	2	4	3	1	4	7	3.44	3	1.63	54
70	5	6	3	4	2	6	6	2	5	4.33	4	1.60	40
60	7	7	7	6	6	7	5	7	1	5.89	5	2.05	41
50	1	4	6	7	3	2	8	5	3	4.33	6	2.77	46
40	6	5	5	5	7	5	3	8	4	5.33	7	2.19	31
30	8	8	9	8	9	8	7	6	8	7.89	8	0.80	10
20	10	9	8	10	8	9	9	9	9	9.00	9	0.66	7
10	9	10	10	9	10	10	10	10	10	9.78	10	0.47	6
0	11	11	11	11	11	11	11	11	11	11.00	11	0	0

'correct' position; these quantities are shown in the last two columns. Because of the very high *SD* for the 15% solution, it was thought advisable to make another series of tests. The results of these, five runs each by two *O*s, are shown in Table VIII.

The results of similar experiments made with solutions of acetone in water are shown in Table IX and of others with solutions of ethyl mercaptan in Table X.

DISCUSSION

(1) *Agreement among the O*s. Examination of Tables I and II shows that the responses of the three *O*s to solutions of acetone after adaptation

with pure acetone or with water are very similar; the PR_{50} value for each O is usually, although not always, the same as that of the group. When variations do occur they are not very large; for example, in Table I, the PR_{50} concentration for SL and RMM and for the group is 5% although

TABLE X

ORGANOLEPTIC APPRAISAL OF SMELL INTENSITY OF DILUTIONS OF ETHYL MERCAPTAN

Concentration ethyl mercaptan in water (% volume)	Positions assigned to coded solutions								Group mean	'Correct' position	SD from 'correct' position	Coefficient of variation (%) from 'correct' position
	RWM				SL							
	1st run	2nd run	3rd run	4th run	1st run	2nd run	3rd run	4th run				
1.0	1	1	1	1	1	1	1	1	1.00	1	0	0
0.3	2	2	3	2	2	2	2	2	2.13	2	0.35	18
0.1	3	3	2	3	3	3	3	3	2.88	3	0.35	12
0.03	4	4	4	4	4	4	4	4	4.00	4	0	0
0.01	5	5	5	5	5	5	5	5	5.00	5	0	0
0.003	7	6	6	6	6	6	6	6	6.13	6	0.35	6
0.001	6	7	7	7	7	7	7	7	6.88	7	0.35	5
0.0003	8	8	8	8	8	8	8	8	8.00	8	0	0
0.0001	9	9	9	9	9	9	9	9	9.00	9	0	0
0.00003	10	10	10	11	11	10	10	10	10.25	10	0.50	5
0.00001	11	11	12	10	10	12	11	11	11.00	11	0.71	6
0.0	12	12	11	12	12	11	12	12	12.00	12	0.50	4

that for RWM is 6%. The agreement was reasonable, too, on the other odorants that were tested and for which the primary observations are not recorded in detail; there were no enormous disparities.

(2) *Odor-intensity of pure odorants.* All of the substances used in this part of the investigation were undiluted except for ethyl mercaptan; it is physically impossible to take deliberate sniffs of this substance unless it is greatly diluted, and accordingly a 0.5% solution of ethyl mercaptan in water was included amongst the undiluted odorants.

These odorants are arranged in alphabetical order in Table III; if they were arranged in order of odor-intensity we should find first place taken by the 0.5% solution of ethyl mercaptan in water with an intensity of 14,000; as will be shown later the intensity of undiluted ethyl mercaptan may be about 200,000. It is well known that mercaptans have most intense odors and it is not unexpected to find the representative from this group of compounds at the top of the intensity tree. Second place is taken by allyl caproate (5000) which has a most intense and irritating smell of pineapples (the irritant action is doubtless due to the chemical unsaturation in the allyl group). Next comes pyridine (1400); pyridine has an intense odor but one not nearly so intense as those of the mercaptans. Then come those substances which have strong, but not unbearably intense, smells, such as *n*-butanol, oil of cloves, acetone, amyl acetate, butyric acid and diacetone alcohol. Eventually we reach those substances with weak odors: methanol has an odor-intensity of only 1.0 and benzylamine one of only 1.7. All of the odorants fall into acceptable places; the most intense have high values and the weakest have the lowest.

This empirical agreement of observed and calculated intensities supports the view that the method of deriving values of intensity is soundly based. Furthermore, every value of intensity is obtained from physical measurements of threshold concentrations and of self-adaptation, two factors which are evidently related to intensity of odor. There seems to be some justification for putting forward a definition of odor-intensity based on these considerations and on the experimental findings, and this, it is suggested, might read thus:

(3) *Definition of odor-intensity.* The intensity of a stimulus-odor is

TABLE XI
ODOR-INTENSITIES OF ISOMERIC BUTANOLS

<i>n</i> -Butanol	$\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_2\text{OH}$	200
<i>sec</i> -Butanol	$\text{CH}_3\text{CH}_2\text{CHOHCH}_3$	40
	CH_3	
<i>iso</i> -Butanol	CHCH_2OH	15
	/	
	CH_3	
	CH_3	
	\	
<i>tert</i> -Butanol	C	20
	/ \	
	CH_3 OH	

the number of times its normal threshold concentration is enhanced by an immediately prior inspiration of the undiluted odorant. An alternative expression is: The odor-intensity of a substance is the ratio of the concentration of the substance's odor-threshold, after one just prior inspiration of the undiluted odorant, to the concentration of the normal, unadapted odor-threshold.

(4) *Odor-intensities of isomers.* The 22 substances listed in Table III include a number of structural isomers; their inclusion was purposeful as it enables a comparison to be made of different chemical entities in which differences of chemical constitution, one from another, are at a minimum. The results that these isomers gave are somewhat surprising: whereas *n*-propanol has an odor-intensity of only 14, *iso*propanol has a value of 28, the branched chain being associated with a slightly more intense odor. Yet of the four butanols which had values as shown in Table XI, the straight chain compounds had the most intense odors.

(5) *Odor-intensities of homologues.* The straight chain aliphatic alcohols afford the comparison shown in Table XII. Examination of this table shows that there is an increase in odor-intensity with each rise in the homologous series; and, indeed, if one smells the first four straight chain alcohols there is no difficulty in deciding that their odor-intensities are in the order butanol > propanol > ethanol > methanol. The surprising feature of

Table XII is the enormously greater odor-intensity of *n*-butanol than of the lower members of the series; the jump from *n*-propanol to *n*-butanol is very large. Yet if one compares the smells of these two alcohols there is not the least doubt that the odor of *n*-butanol is indeed overpoweringly strong compared with that of *n*-propanol. The 'anomaly' in the intensity

TABLE XII
ODOR-INTENSITIES OF HOMOLOGOUS ALCOHOLS

Methanol	CH ₃ OH	10
Ethanol	CH ₃ CH ₂ OH	12
<i>n</i> -Propanol	CH ₃ CH ₂ CH ₂ OH	14
<i>n</i> -Butanol	CH ₃ CH ₂ CH ₂ CH ₂ OH	200

figures is only a reflection of the surprisingly strong smell that *n*-butanol really has.

(6) *Odor-intensity of dilutions of odorants: (a) Acetone.* Intensity of odor quite clearly diminishes with increasing dilution; whereas undiluted acetone has an overpowering odor, a 1% solution in water has only a mild odor and a 0.1% is only just detectable.

The concentration used for adaptation in any particular experiment, *e.g.* 10% acetone, is considered for the purpose of this investigation as a pure (undiluted) odorant. The intensity of the adapting odorant, *e.g.* 10% acetone, is calculated as before; it is unnecessary to make any correction in the calculation for the fact that the acetone is only 10%, because *both* of the threshold concentrations (after acetone and after water) would be multiplied by the same factor, hence the intensity-value would be unaffected; consequently all threshold concentrations can be expressed as of pure acetone.

To make this perfectly clear the case of 10% acetone may be considered. As Table IV shows, the PR_{50} concentration, C_1 , after adaptation with undiluted acetone was 1.5% acetone, or 15% of the 10% acetone; the PR_{50} concentration, C_2 , after 'adaptation' with water, was 0.03% acetone, or 0.3% of the 10% acetone. If the intensity of the 10% acetone is calculated on the basis that the 10% acetone is a 'pure' odorant, we have: $I = C_1/C_2 = 15/0.3 = 50$. On the other hand, if the intensity of the 10% acetone is calculated using threshold concentrations on a pure (100%) acetone basis we have: $I = C_1/C_2 = 1.5/0.03 = 50$. Consequently in Tables IV-VI, the PR_{50} threshold concentrations are expressed as concentrations of the undiluted odorant, whether acetone, *isopropanol*, or ethyl mercaptan. This device simplifies the understanding of the tables and does not affect the intensity figures.

It can be seen from Table IV that the odor-intensity of the acetone solutions rises until a concentration of about 50% is reached and that

thereafter the odor-intensity is practically constant. The high value for 80% may be due to experimental error.

The rise of intensity to a *limiting* value is very evident when the logarithm of the odor-intensity is plotted against the adapting concentration as has been done in Fig. 1. It can also be seen, when the results are plotted as in Fig. 2 that for concentrations up to 50% the logarithm of the odor-intensity is proportional to the logarithm of the adapting

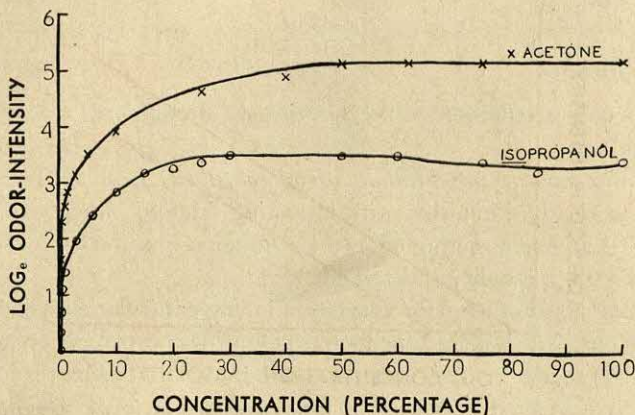


FIG. 1. LOG OF ODOR-INTENSITY VS. STIMULUS-CONCENTRATION

concentration or that: $\log_e I = k(\log_e c - \log_e t)$, where I is the odor-intensity, k is a constant, c is the concentration of the acetone solution, and t is the threshold concentration of acetone after 'smelling' water.

(b) *Isopropanol*. The results obtained with *isopropanol* which are shown in Table V show that once again the odor-intensities of the solutions rise as can also be seen from Fig. 1 until a certain concentration—this time of about 25%—is reached and thereafter the odor-intensity is more or less unchanged; the slight falling off in intensity at the highest concentrations is not understood, it was checked repeatedly and is believed not to be due to experimental error; it is thought that it may be due to a drying action of the concentrated solutions of *isopropanol* on the receptors. But below the limiting concentration of 25% once again odor-intensity progresses geometrically with the geometric progression of the concentration of *isopropanol*; this can be easily seen from Fig. 2 which includes the intensity-curve for *isopropanol* as well as for acetone.

(c) *Ethyl mercaptan*. In any work on smell it is tempting to use one of the mercaptans, and in a study of the intensity of smell it seemed obligatory to do so. There is, however, nothing enjoyable about smelling strong solutions of mercaptans, and anything over 1% is physically impossible.

The group were able to smell 0.3% as an adapting odorant and one member of the group smelt as high as 0.5%. Tests were made also on lower concentrations and the results obtained were as shown in Table VI. The experimental results are plotted on a log/log basis in Fig. 3 and the linear relationship is again evident.

(7) *A limiting intensity.* The concept of a limiting odor-intensity is new, so far as is known. The curves in Fig. 1, and those in Fig. 2, although

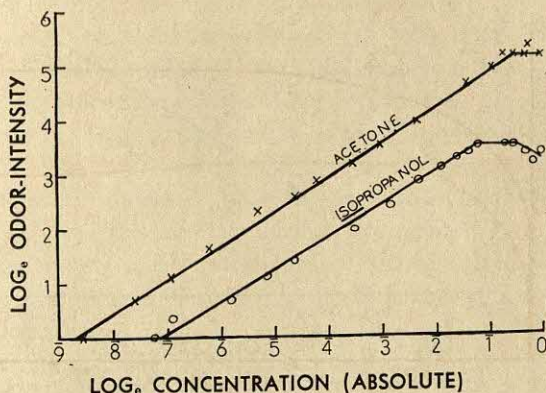


FIG. 2. VARIATION OF THE LOG OF THE ODOR-INTENSITY OF ACETONE AND ISOPROPANOL LOG OF THE STIMULUS-CONCENTRATION

not quite so clearly, show that the odor-intensity rises steeply at first with increase of concentration, but then falls off and remains more or less stationary over a wide range of concentrations.

It does not seem difficult to interpret this behavior from physiological considerations. The concentration of odorant molecules in the air above a solution of an odorant will increase arithmetically with the concentration of odorant in the solution. Evidence has recently been put forward that the primary stimulus for the sensation of smell is a process of *adsorption* of the odorant molecules on the olfactory receptor system and as the concentration of odorant molecules in the inspired air increases, the number of adsorbed molecules at any moment will increase and the intensity of the smell sensation will increase.¹¹ This process cannot go on indefinitely, however, because the area of the olfactory receptors is relatively small and a stage may be reached at which all those receptors which will accommodate acetone molecules for example, are filled—a monomolecular layer of odorant molecules completely covers them. When this stage is reached a further increase in the concentration of the odorant molecules will have no effect; no more can be adsorbed than are already adsorbed and conse-

¹¹ R. W. Moncreiff, The characterization of odours, *J. Physiol.*, 125, 1954, 453-465; The sorptive properties of the olfactory membrane, *ibid.*, 130, 1955, 543-558.

quently there will be no increase in intensity of odor. This is exactly what was found to happen in the case of acetone; the intensity of the odor increased steadily until a concentration of 50% was reached. Similar considerations apply to *isopropanol* which reaches a maximal intensity at a concentration of about 30%.

With the relatively low concentrations of mercaptan that were used the saturation point of adsorption on the receptors was not reached. The

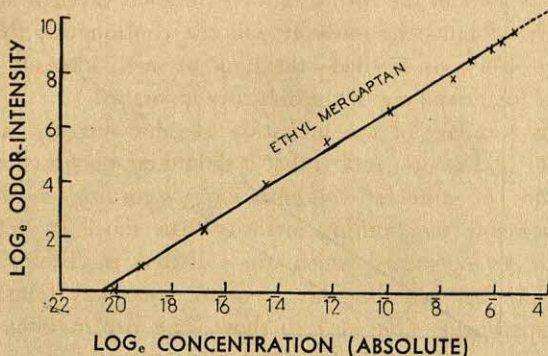


FIG. 3. VARIATION OF THE LOG OF THE ODOR-INTENSITY OF ETHYL MERCAPTAN WITH THE LOG OF THE STIMULUS-CONCENTRATION

relation between the logarithm of odor-intensity and the logarithm of the concentration still holds very closely as can be seen from Fig. 3.

It is interesting to note that if saturation of the receptors were not reached earlier (as it probably would be) the odor-intensity of undiluted ethyl mercaptan would be about 200,000, something of the order of 1000 times the intensity of acetone. Those who have smelt strong mercaptan solutions will have no difficulty in accepting such a figure.

The limitation of intensity in dilutions of a powerful odorant is, it must be emphasized, imposed by man's olfactory apparatus and not by the odorant material. When solutions of 50, 60, 70, 80 and 90% acetone are smelt directly from a bottle they smell much the same, because the lowest of the concentrations provides sufficient odorant molecules to complete an adsorbed layer on the appropriate receptors. If, however, the solutions were smelled at a distance then the lowest would not provide sufficient odorant molecules in the inspired air to complete an adsorbed layer and the next highest would provide more molecules so that the adsorbed layer on the receptors would have a greater area. Thus 90% acetone smelled with the bottle a few inches away from the nose is obviously stronger than 50%; it is only when conditions are such that the lowest concentration can afford sufficient odorant molecules to fill the appropriate receptors, that a

limitation of intensity is reached. Clearly, a smell, just like a source of light or a sound, becomes less intense as its distance from *O* is increased. It is, however, possible to have conditions under which diluted acetone or *isopropanol*, and doubtless most other strong odorants, will give the maximal physiological response, so that thereafter no increase in concentration is detectable as an increase in intensity. In the case of weak odorants the olfactory apparatus is capable of adsorbing the molecules in the air above it without the formation of a complete layer of adsorbate, so that with such odorants the intensity will rise continuously through their dilutions until 100% is reached—the limit of smell intensity is thus imposed by the odorant and not by the olfactory apparatus.

An obvious corollary of the limitation of odor-intensity in substances such as acetone and *isopropanol* is that it should be much easier to arrange coded solutions of different concentrations, according to their correct concentration below the limiting intensity than above it. When the results of such arrangements which are collected in Tables VII-X are examined it can be seen that this is so. Acetone was found from the adaptation experiments (Fig. 1.) to show little rise in intensity beyond 40% concentration. Table IX shows that the standard deviation of placements of concentrations from the 'correct' value is about 2 for concentrations of 40% and over but is less than 1 for concentrations below 40%; the accompanying drop in the coefficient of variation from about 40 to 10 or less is even more striking; far more mistakes in order of arrangement are made above than below 40% because the intensities are much more similar above 40% than below this concentration.

The log intensity concentration curve (Fig. 1) for *isopropanol* shows a marked falling off in rate of increase of intensity with concentration above 20%, the intensity being determined by experiments on adaptation. The sorting of the various concentrations into order of concentrations can be seen from Table VIII to show a sudden improvement in accuracy below 20%. Whereas from 20% to 90% *isopropanol* the standard deviation from the 'correct' placement is of the order of 3.5 to 4, it drops suddenly to 1.67 for 15% and then to even lower values. The figures for *isopropanol* in Table VII show very similar characteristics, apart from the anomalously high standard deviation of the 15% solution; this value is apparently due to chance as there is no sign of its re-appearance in Table VIII.

Ethyl mercaptan as can be seen from Fig. 3 shows no approach to a limiting value of odor-intensity within the range of concentrations, up to the maximum of 0.5%, that we found it possible to sniff adequately. Similarly the arrangement of concentrations of ethyl mercaptan shown in

Table X shows no great disparities; the standard deviation from the 'correct' placement never reaches a value of 1; no limit of intensity was reached and there was no difficulty in sorting out the different concentrations according to smell intensity.

(8) *Analysis of the curves.* It can be seen from Figs. 2 and 3 that when the logarithm of the odor-intensity is plotted against the logarithm of the concentration, then a straight line results up to the limiting intensity, irrespective of whether the odorant is acetone, *isopropanol*, or ethyl mercaptan. Of course, if logarithms to the base 10 are used instead of to the base e then the relationships are still straight line.

For all three odorants, the intensity-concentration relationship can be expressed: $\log_e I = k(\log_e c - \log_e t)$ where I is the intensity of odor, c is the concentration, k is a constant, and t is the normal threshold concentration (after having 'smelt' water). The constant k seems to be remarkably uniform; the angles of slope for the three curves are: acetone 33° (tangent = 0.65); *isopropanol* 31° (tangent = 0.60); ethyl mercaptan 32° (tangent = 0.62). Cheesman and Mayne (1953) found that very similar curves (adaptation in homogeneous pairs) for *isopropanol*, dioxan, and *cyclopentanone* had a common gradient of 0.7. Our own experimental results point to a figure of 0.62 or the vulgar fraction $5/8$.

The values of $\log_e t$ are for acetone $\bar{9}.9$, for *isopropanol* $\bar{8}.7$, and for ethyl mercaptan $\bar{20}.0$. The relation between the intensity and the concentration for all three odorants may therefore be expressed: $\log_e I = 5/8(\log_e c - \log_e t)$. Another relationship may be derived from this one as follows:

By definition

$$I = t_c/t,$$

where t_c is the threshold concentration after prior smelling of concentration c and t is the threshold concentration after smelling water or $\log_e I = \log_e t_c - \log_e t$. Hence:

$$\begin{aligned}\log_e t_c - \log_e t &= 5/8(\log_e c - \log_e t) \\ &= 5/8\log_e c - 5/8\log_e t \\ \log_e t_c &= 5/8\log_e c + 3/8\log_e t \\ \text{or } t_c &= c^{3/8}t^{3/8}\end{aligned}$$

Knowing the threshold concentration of an odorant it is possible from this relationship to calculate what the threshold concentration of the odorant will be after prior smelling of any given concentration of it. This relationship applies until the limiting intensity of the odor is approached; when this happens t_c no longer rises with c .

It is interesting to note that the basic relationship

$$\log_e I = k(\log_e c - \log_e t)$$

in which k has a value of 0.62 is not very far removed from a square root relationship. Thus if the odor-intensity had been proportional to the square root of the concentration of an odorant, the value of k would have been 0.5. In fact the value of k is about 0.62. It is nevertheless a useful approximation, to remember that odor-intensity is roughly proportional to the square root of concentration.

The influence of a low threshold concentration on intensity of odor is clearly seen in Fig. 4 in which part of the curve, only just determined,

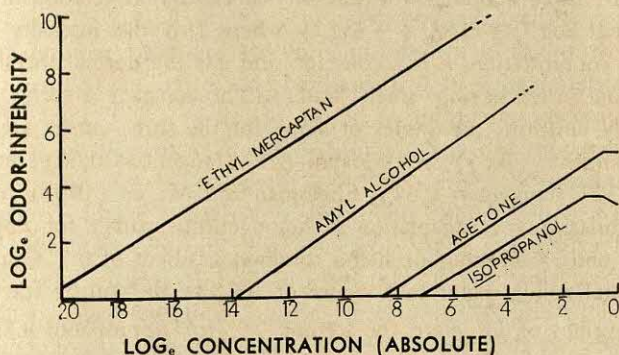


FIG. 4. DEPENDENCY OF ODOR-INTENSITY ON THRESHOLD CONCENTRATION

for amyl alcohol (B.S.696) has been included; the very low threshold concentration of ethyl mercaptan is the cause of the extreme intensity of this substance in even quite low concentrations. With substances such as mercaptans the olfactory receptors can be stimulated by a relatively small number of molecules; with other substances such as isopropanol that have much higher threshold concentrations, a much greater number of molecules is necessary to stimulate the receptors. Once the number necessary to stimulate is exceeded, whether for mercaptan or acetone or any other odorant, the intensity of smell sensation rises a little more but not much more rapidly than the square root of the concentration of odorant. Eventually those receptors that adsorb the molecules of a particular odorant become nearly completely covered with a layer of adsorbate and then the intensity rises more and more slowly and finally reaches a limiting value. Two physical measurements must necessarily be made on any odorant in order to be able to predict the intensity of any of its dilutions; they are: (1) the threshold concentration without previous adaptation; and (2) the limiting intensity. The first would seem to be determined by the sensitivity

of the receptors appropriate to that odorant and the second in part by their abundance and in part by the time of adsorption of the molecules of a particular odorant on the adsorbent surface—the longer this time, perhaps 10^{-4} to 10^{-2} sec., the greater will be the coverage of adsorbent receptors by adsorbate odorant molecules.

SUMMARY

(1) Odor-intensity is considered as a function of (a) a low olfactory threshold concentration and (b) a powerful adapting effect.

(2) The odor-intensity of a substance is defined as the number of times that its normal odor-threshold concentration is enhanced by one just prior inspiration of the undiluted substance.

(3) Assessment of the odor-intensity of a substance requires the experimental determination of two threshold concentrations. *O* gives only 'yes' or 'no' responses according as he can or can not smell a given solution; he is not required to give any other opinions or to *compare* smells, either like or unlike.

(4) The scale-values of odor-intensity arrived at are in accordance with practical experience; those substances with the most intense odors have the highest values, *e.g.* 14,000 for a 0.5% mercaptan solution, 1400 for undiluted pyridine, 150 for amyl acetate, while substances with weaker odors have much lower values, *e.g.*, 28 for *isopropanol*, and 10 for methanol.

(5) The intensity of solutions of odorous substances in odorless diluents rises with increase in concentration up to a limit. With acetone solutions smelled directly from a wide-neck bottle this limit is reached at about 40%, with *isopropanol* at about 20%. The increase in odor-intensity is believed to run parallel with increased adsorption of the odorant molecules on the receptor surfaces; so long as this can increase, odor-intensity can increase, but when once the receptor sites that will provide lodgement for acetone molecules for example, are filled with an adsorbed monomolecular layer, no further increase in intensity is possible.

(6) Up to the limiting value, the odor-intensity of solutions increases geometrically with the geometric progression of their concentration. This relationship appears to hold well right from the threshold concentration to the limiting intensity for acetone and for *isopropanol* and for solutions of ethyl mercaptan as strong as it is physically possible to smell them.

(7) As a rough approximation, the odor-intensity of solutions of an odorant is proportional to the square root of the concentration of the solution.



SOME DETERMINANTS OF PHENOMENAL OVERLAPPING

By DOROTHY DINNERSTEIN, Bank Street College of Education, and
MICHAEL WERTHEIMER, University of Colorado

The experiments reported here were designed to test several aspects of a general formulation concerning the phenomenal overlapping of visual planes; namely, that overlapping depends on such gestalt principles as symmetry, simplicity, and similarity;¹ and that the structure whose organization determines overlapping in a particular instance embraces not only the overlapping areas themselves but also wider portions of visual space.

Since these experiments were conducted, several discussions of this problem have been published. Ratoosh, following Holmholtz, has offered a formulation in terms of local continuity of contours.² Gibson stresses not only continuity but also "completeness" and "regularity" of outline, at the same time suggesting that the effectiveness of these factors may depend on learning.³ Chapanis and McCleary, in a paper contesting the adequacy of Ratoosh's formulation, present evidence also pointing in a general way to the role of gestalt factors and of familiarity.⁴ Arnheim, taking up this problem as part of a comprehensive discussion of visual organization, presents a viewpoint essentially identical with ours.⁵ The present study casts additional light on the limitations of the Ratoosh-Helmholtz hypothesis; goes somewhat further than the Chapanis-McCleary study into the nature of the gestalt factors determining overlapping; and supports an analysis parallel to Arnheim's with empirical data of a type not offered by him.

We started with two questions. (1) Why is a contour—i.e., a line marking off a visual area—usually one-sided, 'belonging to' and functioning as the edge of only one of the two areas which touch it? (2) What determines to which side the contour will belong? The lines of Fig. 1, for instance, are seen most easily as belonging to the triangular area which

* Received for publication December 8, 1955. This study was made at Swarthmore College in 1946. Our thanks are due Dr. Hans Wallach for his suggestions and assistance.

¹ Max Wertheimer, *Untersuchungen zur Lehre von der Gestalt: II, Psychol. Forsch.*, 4, 1923, 301-350.

² Philburn Ratoosh, On interposition as a cue for the perception of distance, *Proc. Nat'l. Acad. Sci.*, 35, 1949, 257-259.

³ J. J. Gibson, *The Perception of the Visual World*, 1950, 142-143.

⁴ A. Chapanis and R. A. McCleary, Interposition as a cue for the perception of relative distance, *J. Gen. Psych.*, 48, 1953, 113-133.

⁵ Rudolph Arnheim, *Art and Visual Perception*, 1954, 198-203. We did not know of Dr. Arnheim's prior work in this area until after completing our experiments. Our study adds little, in the way of theory, to the formulation which he has since published.

they enclose, with the adjacent surface continuing underneath. Why is it harder to see them as shared by both areas—as if a triangular hole were cut in the paper and a triangular plug fitted into it? If they must be one-sided, why do they not belong to the other side—as if a triangular hole were cut in the paper through which we viewed another surface of undefined boundary?

In the case of Fig. 1, it might seem plausible to answer these questions in figure-ground terms, explaining the ownership of the contour as a consequence of the fact that the triangle has greater object-quality than

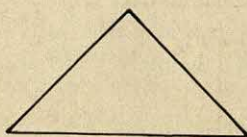


FIG. 1. ENCLOSED TRIANGULAR AREA

the area continuing underneath.⁶ Later, we shall show that the 'thing-stuff' distinction cannot always explain phenomenal overlapping even when the areas in question do differ in this respect. But meanwhile, we shall begin our inquiry with some cases in which *both* the areas separated by the contour are 'figures,' or 'things.'

GENERAL PROCEDURE

India ink drawings, on white cards—Figs. 2-18 were about 11×15 in., Figs. 19 and 20 about 8×10.5 in.—were shown to assembled groups of students, who wrote descriptions of what they saw. Our instructions were: "Describe what this is. State what surfaces, if any, appear to go behind other surfaces." Each card was held up, for about 15 sec., until everyone in the group appeared satisfied that he had seen it and then about 2 min. were allowed for writing before the next card was shown.

For analysis, these descriptions were later categorized in terms of contour-belongingness. This was generally, as may be seen from some of the samples quoted below, a fairly simple matter. Where the experimenters felt uncertain, or disagreed, after reading a description of a figure, about how a particular *S* had assigned its contours, the response was categorized as 'ambiguous.' In the rare cases in which *S* described two alternate organizations of a figure, the response was categorized on the basis of the first description.

Every group of *Ss* was shown a number of drawings. We did not try to control the influence of order of presentation, *i.e.* the influence of relatively unrelated dissimilar drawings on each other; such influences, while undoubtedly present, seem very unlikely to have been of an order which could bear crucially on the results to be reported. Detailed documentation of this point would require inordinate

⁶ Edgar Rubin, *Visuell wahrgenommene Figuren*, 1921.

space, since many of the drawings shown are not discussed here. In some cases, however, where we wished to compare responses to closely related drawings, we presented them to separate groups of Ss. We indicate below for each comparison whether it is based on the same group of Ss (in which case we indicate the order of presentation of the drawings discussed) or on separate groups.

EXPERIMENTS AND RESULTS

(1) *Adjoining visual areas.* Figs. 2-5 were shown in order, to a group of 21 Ss. Table I summarizes the distribution of types of responses.

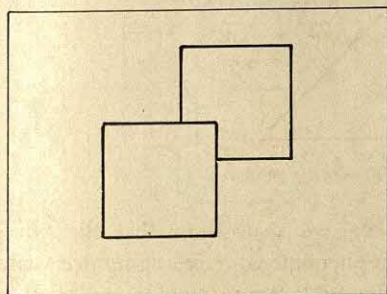


FIG. 2. OVERLAPPING SQUARES

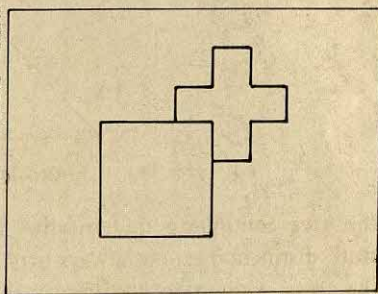


FIG. 3. SQUARE AND CROSS

Throughout, we are foregoing the usual statistical tests of significance, since with a few minor exceptions the differences obtained were so large as to be clearly reliable.

For Fig. 2, our Ss unanimously assigned the crucial contour to the lower left (southwest) area. They reported "overlapping squares," or, more accurately, "two squares so overlapping that the lower left-hand corner of the top one is covered by the upper right-hand corner of the lower one." According to our view of the problem, this assignment of the contour serves to allow the northeast area to achieve symmetry by continuing underneath. This view is tested by Fig. 3, which is the same as Fig. 2 except that the upper and right corners of the northeast area have been cut out. Now this area is symmetrical when seen to end at the common boundary. Here 90% of our Ss assigned the contour equally to both areas, in such terms as "square and cross, touching but not overlapping," or "square and cross—square lines are also two lines of cross."

In Fig. 4, this same cross is moved slightly; now it is the square which can achieve symmetry only by continuing underneath the common border. Again, for 90% of our Ss the contour accordingly belongs to the cross. The figure is described, for example, as "cross overlapping square," or

"the cross is partly superimposed on the right-hand upper corner of the square."

Thus far our results may be explained entirely in terms of the shape of the area which continues underneath. One might conclude: "Whenever an area can improve its shape by doing so, it will appear to extend behind the adjacent area." Fig. 5 shows that this explanation is incomplete. Here the northeast area is restored to the shape it had in Fig. 2, which

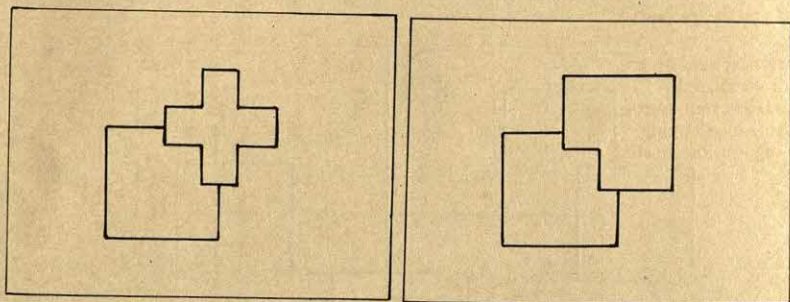


FIG. 4. CROSS OVERLAPPING SQUARE

FIG. 5. HEXAGON OVERLAPPING SQUARE

in that situation it improved for 100% of our Ss by yielding the contour to its neighbor and continuing underneath. Now, combined with a neighbor which competes for the opportunity to continue underneath, it ends at the contour for 62% of our Ss. These Ss describe the experience as "one complete square, covered at upper right corner by a second square, which has square removed from lower left corner"; or "square and six-sided figure (indented square) which overlaps first square." Thus, contour-belongingness cannot be explained exclusively in terms either of the area in front (see Figs. 2 vs. 3, and 4 vs. 3) or of the area behind (see Figs. 2 vs. 5). It is a function of the relation between these areas, of the larger organization which together they form.

Fig. 5 can also be seen as two complete squares, the lower left corner of one being inserted in a right-angled slot near the upper right corner of the other. This was probably (note question mark on Table I) the organization seen by the Ss who described Fig. 5 simply as "two squares." In any case, the reader will note that this interpenetrating organization of the figure is possible, and that it carries with it a feeling of satisfaction.

This feeling is even stronger in connection with Fig. 6, when the conglomerate of polygons and triangles is resolved into three interlocked

squares. The greater satisfaction may be related to the greater effort required to see this organization, and to the greater contrast between its snug unity and the incoherence which it relieves. Table II shows that half of the *Ss* who saw Fig. 6 did thus resolve it into the fewest possible units, describing three squares "jutting through," "inserted through slits of,"

TABLE I
DISTRIBUTION IN PERCENTAGE OF TYPES OF DESCRIPTIONS OF FIGS. 2, 3, 4, AND 5
($N=21$)

Type of description	Figs.			
	2	3	4	5
Southwest area on top	100	0	0	0
Contour shared	0	90	5	10
Northeast area on top	0	5	90	62
Mutual overlapping (?)	0	0	0	10
Ambiguous or omitted	0	5	5	18

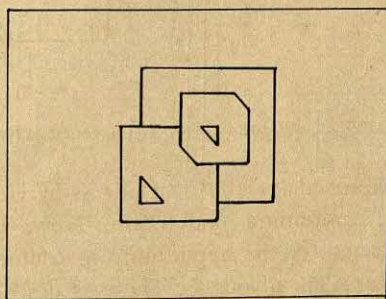


FIG. 6. THREE OVERLAPPING,
INTERPENETRATING SQUARES

TABLE II
DISTRIBUTION OF TYPES OF DESCRIPTIONS OF FIG. 6
($N=21$)

Type of description	%
Three squares interlocked	52
Three squares and triangles	59
Three polygons and triangles	14
Ambiguous	5

"interlocking with," or "merging with" each other. Other *Ss* saw squares and triangles, perhaps a compromise embodying both the tendency of the larger areas to complete themselves and the tendency of the smaller areas to remain within their compact and locally stable limits. Finally, a few described irregular polygons and triangles.

The 'merging' or 'interlocking' organizations of Figs. 5 and 6 make clear the inadequacy of the figure-ground distinction as an explanation of overlapping. In our earlier examples, it might have been argued that the area behind served, relatively speaking, as ground for the area in front. This argument, in the instances of Figs. 5 and 6 would have to regard a given area as serving simultaneously as figure and ground in relation to its neighbor. It seems more plausible to regard the classical figure-ground situation as a special case of phenomenal overlapping. We return to this problem in a later section.

Meanwhile, Fig. 6 requires in another connection that we extend the formulation which fitted the previous four figures. In this case the belongingness of certain crucial contours depends, not only on the con-

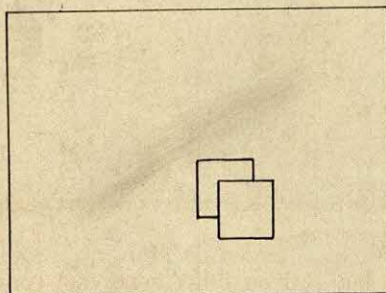


FIG. 7. OVERLAPPING SQUARES

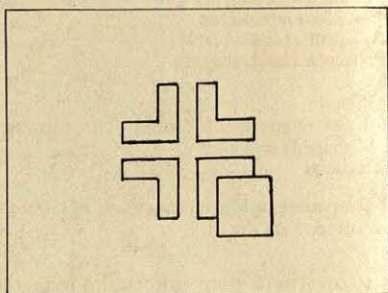


FIG. 8. SQUARE AND FOUR
'L'-SHAPED AREAS

tiguous areas themselves, but also on some additional, non-contiguous, areas. Thus, for those Ss (52%) who saw the southwest triangle in Fig. 6 as a corner of the largest square, the diagonal line of this triangle belonged to—functioned as a slit in—the area to its northeast. This assignment of the contour cannot be accounted for by the shapes of the two areas which it touches. Left to themselves, these would surely overlap the other way, the triangle ending at its borders and claiming all three as contours, and its companion continuing unbroken behind. It is the organization of the outlying areas of Fig. 6 which determines this local outcome.

(2) *Outlying visual areas.* Figs. 7 and 8 provide a sharper test of this extended formulation. Table III summarizes responses to these figures, which were shown to separate groups of Ss. Fig. 7 (like Fig. 2, which it closely resembles) is seen almost always as two overlapping squares. In Fig. 8, this pair of areas reappears, accompanied by a constellation of

areas whose shapes are identical with the 'L' which the northwest member of the pair becomes if it ends at the common border and asserts an equal claim on the contour. The presence of this constellation reduces from 97% to 57% the proportion of Ss for whom the crucial contour belongs exclusively to the southeast member of the pair, and correspondingly raises from 3% to 34% the proportion for whom it belongs equally to both members. Thus, the overlapping of these two areas depends not

TABLE III
DISTRIBUTION IN PERCENTAGE OF TYPES OF DESCRIPTIONS OF FIGS. 7 AND 8
($N=64$ and 56 respectively)

Type of description	Figs.	
	7	8*
Southeast square on top:		
Adjacent area a square	97	30
Adjacent area L-shaped	0	27
Total	97	57
Southeast square shares contour with adjacent ('L'-shaped) area	3	34
Ambiguous	0	9

* Responses to Fig. 8 categorized in terms of how S describes portion of this figure which is identical with Fig. 7.

only on their own internal character, but also on other areas, not contiguous with the crucial contour.

For about half of those Ss who assigned the crucial contour in Fig. 8 to the southeast area, there occurred a compromise. The northwest member of the pair does continue behind its neighbor, but as an 'L' rather than as a square. Here the overlapping cannot be explained as allowing the area behind to 'complete' or to 'improve' itself, since by virtue of its membership in the constellation it remains an 'L' even under partial cover. The one-sidedness of the contour occurs, furthermore, at the expense of the perfect symmetry of the constellation, since to continue behind the contour this 'L' must become somewhat thicker than the others. (Our protocols do not tell us whether this asymmetry was perceived or blurred over.) This seems to be an illustration of the arbitrary one-sidedness of the contour function, a problem to which we shall return.

Figs. 9 and 10, also shown to separate groups of Ss, provide another test of the influence of outlying visual areas. As Table IV shows, the presence of the constellation of 'L'-shaped areas in Fig. 10 substantially changes the belongingness of the crucial contours of the two central areas. The proportion of Ss who saw these areas as two squares drops from 92% to 28%, while the proportion who saw them as two 'L'-shaped areas rises correspondingly from 3% to 44%. Ambiguous descriptions of Fig. 10

(which is both perceptually unstable and hard to describe) reached a high of 28%.

Responses to Fig. 9 were interesting also in another connection. Among the Ss who described this figure as two squares, there were some who explained that these were "transparent," "wire" or "line" squares. The fact that none of the figures discussed above were ever described in such terms suggests that lines in drawings like this function as line-figures only

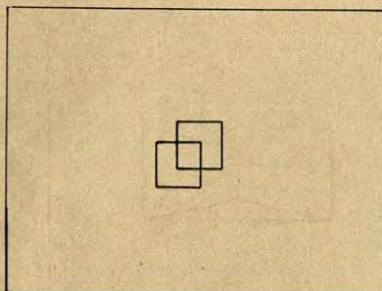


FIG. 9. TWO SQUARES, TRANSPARENT OR WIRE

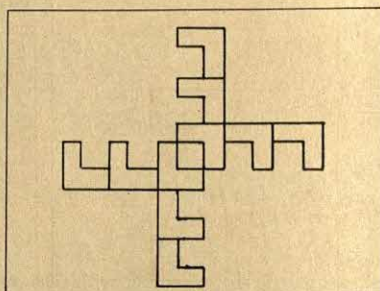


FIG. 10. CONSTELLATION OF 'L'-SHAPED AREAS

TABLE IV

DISTRIBUTION IN PERCENTAGE OF TYPES OF DESCRIPTIONS OF FIGS. 9 AND 10
($N=64$ and 54 respectively)

Type of description	Figs.	
	9	10*
Two squares	92	28
Two 'L's'	3	44
Omitted or ambiguous	5	28

* Responses to Fig. 10 categorized in terms of how S describes portion of this figure which is identical with Fig. 9.

when this is structurally necessary. The preference to see contours rather than line-figures, so far as structure permits, is also suggested by the fact that some Ss described a superimposed square (usually the one to the southwest) which was transparent (or wire or line) while its companion remained opaque. Below, we shall cite further evidence bearing on this point.

(3) *Main directions of visual space.* So far we have considered only the organization of the drawing itself. The effective visual field actually, however, extends further. Figs. 11 and 12 show the same drawing, presented to separate groups of Ss in a vertical and a horizontal position to test the effect, on overlapping, of the relation between the drawing as a whole

and the main directions of visual space. In Fig. 11, the line separating the hexagonal and triangular areas must serve as a contour of the hexagon to preserve the horizontal symmetry of that area, while in Fig. 12 the hexagon's horizontal symmetry is unaffected by the belongingness of the contour, only its vertical symmetry being at stake. As Table V shows, 66% of those viewing Fig. 11, as compared with 29% of those viewing Fig. 12,

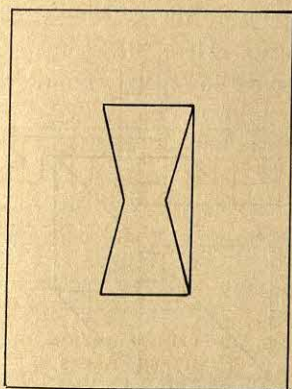


FIG. 11. HEXAGON AND TRIANGLE
(Vertical position)

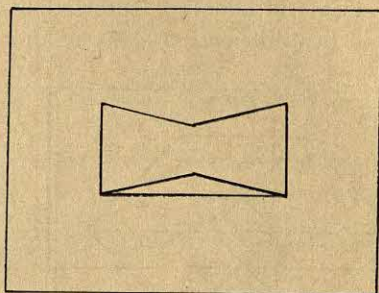


FIG. 12. HEXAGON AND TRIANGLE
(Horizontal position)

TABLE V

DISTRIBUTION IN PERCENTAGE OF TYPES OF DESCRIPTIONS OF FIGS. 11 AND 12
($N=38$ and 42 respectively)

Type of description	Figs.	
	11	12
Hexagonal area on top	66	29
Triangular area on top	7	21
Contour shared	3	2
Tridimensional	5	19
Omitted or ambiguous	18	29

assigned the crucial contour to the hexagon, *i.e.*, an "hour-glass" or "bow-tie." For some of the Ss, the remaining line was the edge of another area which continued under the hexagon; for others, it functioned as a separate line-figure on a continuous ground. Conversely, the crucial contour was assigned more often to the triangle in Fig. 12 than in Fig. 11, although this difference (21% vs. 7%) may not be significant.

Finally, three-dimensionality ('prism,' 'corner of a room,' etc.)—a type of response which we did not anticipate—occurred more frequently for

Fig. 12 than for Fig. 11. Of course it makes sense that the factors determining overlapping—which is simply the flattest kind of three-dimensionality—should also influence more drastic invasions of the depth dimension. This difference (19% versus 5% cent) is again, however, too small for clear significance.

(4) *Other modes of visual organization: Depth, line-figures.* Figs. 13 and 14 test more sharply the suggestion, derived from Figs. 11 and 12, that phenomenal depth is subject, like overlapping, to the drawing's relation to the main directions of surrounding space. Here the same pattern, drawn in two different positions, was presented to separate groups of Ss.

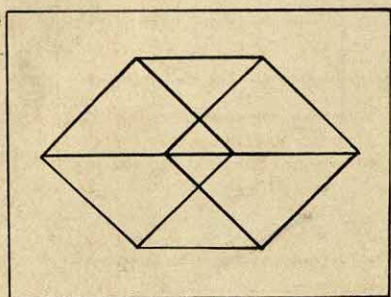


FIG. 13. LINE-PATTERN
(More frequently seen as two-dimensional pattern)

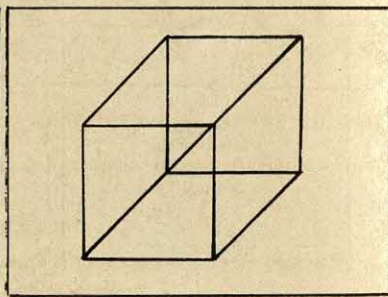


FIG. 14. LINE-PATTERN
(More frequently seen as cube)

TABLE VI

DISTRIBUTION IN PERCENTAGE OF TYPES OF DESCRIPTIONS OF FIGS. 13 AND 14
($N=64$ and 54 respectively)

Type of description	Figs.	
	13	14
Tridimensional: cube	31	96
Bidimensional pattern	63	0
Omitted or ambiguous	6	4

As Table VI shows, Fig. 14 was much more frequently seen as a cube than Fig. 13 (96% vs. 31%). The factors determining the direction of overlapping also seem related to another question of visual organization, the question of whether lines will function as figures in themselves or as the edges of areas. This relation, suggested above in connection with Fig. 9, becomes clearer in connection with the next two figures (Figs. 15 and 16).

Table VII shows the responses of separate groups of Ss to Figs. 15 and

16, in which the same two curves appear in symmetrical and in parallel alignment. Fig. 15 is much likelier than Fig. 16 (85% vs. 26 %) to assume some organization involving overlapping planes. Conversely, Fig. 16 is much likelier than Fig. 15 (45% vs. 4%) to appear as isolated

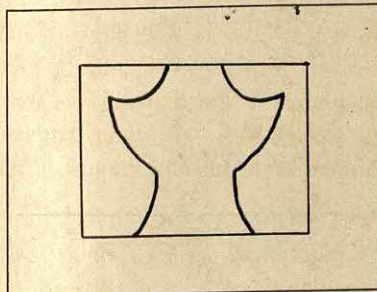


FIG. 15. CURVES IN SYMMETRICAL ALIGNMENT
(More frequently seen as overlapping planes)

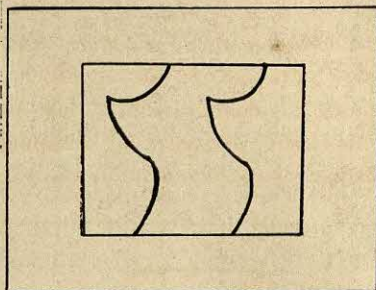


FIG. 16. CURVES IN PARALLEL ALIGNMENT
(More frequently seen as line figures)

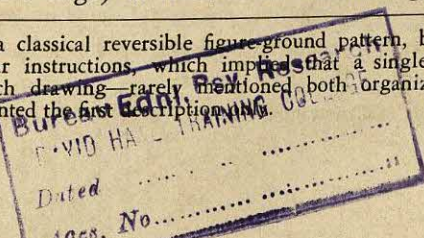
TABLE VII

DISTRIBUTION IN PERCENTAGE OF TYPES OF DESCRIPTIONS OF FIGS. 15 AND 16
($N=54$ and 64 respectively)

Type of description	Figs.	
	15	16
Overlapping planes:		
2 planes, central area on top	63	0
2 planes, side areas on top	22	3
3 planes, successively overlapping	0	23
Total	85	26
Line figures on continuous ground	4	45
Contours shared	0	8
Tridimensional	0	5
Omitted or ambiguous	11	15

lines on a continuous white ground. Some light is cast on this difference by further analysis of the cases in which overlapping did occur. Of those who saw overlapping in Fig. 16, almost all saw *three* successive planes. Of those who saw overlapping in Fig. 15, all saw just *two* planes: usually a central figure (*i.e.* a vase, urn, or shield) in front of a rectangular ground, sometimes a pair of balanced side-figures (*e.g.* curtains at a window or on a stage) in front of an undefined ground.⁷

⁷ Fig. 15 is a classical reversible figure-ground pattern, but our Ss—probably in response to our instructions, which implied that a single description should be offered for each drawing—rarely mentioned both organizations. When this did happen, we counted the first description only.



Clearly, in Fig. 16 the changed position of the curve on the right, impairing the symmetry of the central area and the similarity of the two side areas, reduced the structural advantages and, therefore, the frequency of two-plane organizations. Of the two Ss who did see Fig. 16 as two planes, one, expressing the tension implicit in this view, reported: "Two jagged surfaces, separated by a space, which would fit together if the two were brought together." With a three-decker organization, on the other hand, the parallel curves in Fig. 16 could form the edges of similar figures. Such an organization did accordingly occur for 22% of the Ss who saw Fig. 16; when it possessed no such structural advantage, the organization never occurred.

Apparently, however, the advantages of a (three-decker) overlapping-planes organization for Fig. 16 were slighter than those of a line-figure organization; while the (two-decker) overlapping-planes organization of Fig. 15 was compelling enough to eliminate almost entirely the line-figure alternative. This difference may have to do with some perceptual disadvantage (complexity?) implicit in the piling up of three successive planes.

Of interest, finally, is the small group of Ss for whom Fig. 16 'escaped' into the third dimension. (The reader may notice here again the curious feeling of satisfaction which attended the organization, 'interlocked squares,' of Figs. 5 and 6.) Three-dimensionality was never described for Fig. 15. This difference is certainly too small to be meaningful statistically. Note, however, that for Fig. 15 greater depth than is required for the flattest overlapping of planes would have no structural advantages, while for Fig. 16 the curving forward into space of the central area can correct lateral asymmetry.⁸

Figs. 17 and 18 also bear on the relation between overlapping-plane and line-figure organizations. Table VIII summarizes the responses of the separate groups who saw these figures. In Fig. 17, the diagonal corner line is *always* seen as a contour of the polygon it bounds (which is described as "a square lacking one corner" or a "five-sided geometrical figure"); while in Fig. 18 this same line, now a segment of a longer line, is *never* seen as belonging exclusively to the polygon. The addition of the two segments raises from zero to 40% the proportion of Ss for whom this diagonal functions as a line-figure. It may also at the same time retain some of its contour-quality for these Ss. Their descriptions, "square with a diagonal line cutting off one corner," do not make this

⁸ For a discussion of such gestalt determinants of three-dimensionality, and of the line-figure vs. contour question, see Kurt Koffka, *Principles of Gestalt Psychology*, 1935, 151-164; and Hertha Kopferman, *Psychologische Untersuchungen über die Wirkung zweidimensionaler Darstellung körperlicher Gebilde*, *Psychol. Forsch.*, 13, 1930, 293-364.

clear. The most interesting feature of Fig. 18, however, is its bearing on the figure-ground distinction.

(5) *Distinction between overlapping and figure-ground.* Most Ss saw Fig. 18 as overlapping in an unusual direction: for 60%, the polygon continued underneath the lengthened diagonal, to become a square with its corner hiding behind, viz. a "slit in the paper," "board," or "screen." The polygon's diagonal boundary 'belongs,' in these cases, exclusively

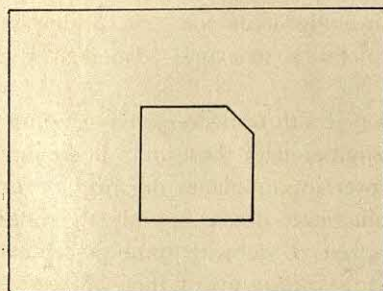


FIG. 17. POLYGON OR MUTILATED SQUARE

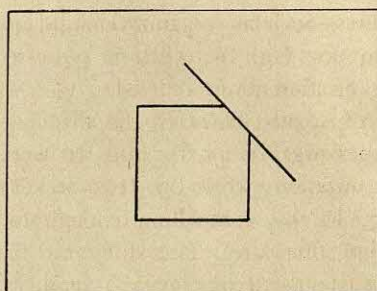


FIG. 18. SQUARE WITH CORNER COVERED

TABLE VIII

DISTRIBUTION IN PERCENTAGE OF TYPES OF DESCRIPTIONS OF FIGS. 17 AND 18
($N=12$ and 42 respectively)

Type of description	Figs.	
	17	18
Polygon or mutilated square	92	0
Diagonal line figure contiguous with polygon or mutilated square	0	40
Square with corner covered by area to northeast	0	60
Omitted	8	0

to the area to its northeast, an area which taken as a while functions as background rather than as figure.

This exceptional case indicates that even when two overlapping planes can be designated as figure and ground—in the sense that one of them has more distinct object-quality than the other—it is not *inevitable* that the contour which separates them will belong to the figure. The fact that the contour *usually* belongs to the figure seems best explained, not in terms of the figure's stronger 'thingness' (see discussion of Figs. 1, 5, and 6), but rather in terms of the ground's tendency to resist interruption, to maintain as much unity and homogeneity as the structure of the field permits. When, as in Fig. 18, other factors outweigh this tendency toward homo-

geneity, the ground can and does 'break,' claim a contour, and phenomenally cover a portion of the figure.⁹ Thus in Fig. 18 the polygon successfully competes with the background for the opportunity to improve its organization by continuing underneath. Clearly this atypical outcome depends on the presence of the additional segments, which—as compared with Fig. 17—weakens the continuity between the diagonal and the adjacent lines. Note, however, that the extended diagonal of Fig. 18 can be seen as a line-figure—and is by 40% of the Ss, and would be by 100% in the absence of the polygon. That it ever functions as a contour of the surround—which thereby loses some unity and homogeneity—must be traced to the presence of the improvable polygon.

(6) *One-sidedness of contours.* One of the questions with which we began was "Why is a contour usually one-sided"? Our findings have on the whole supported the following answer: A contour is usually one-sided because the organization of the field as a whole is usually improved if one of the two areas separated by the contour is seen as continuing behind the other—thereby (if it is background) preserving its homogeneity, or (if it is figure) enhancing either its internal form or its relation with other parts of the field. Which side the contour belongs to depends on which of the two areas can contribute *more* to the organization of the field by continuing behind the other.

When neither area 'demands,' in this sense, to continue behind the other, the tendency toward one-sidedness is very sharply reduced (see, Figs. 2 vs. 3, and 7 vs. 8). It is not, however, completely eradicated. This problem, mentioned above in connection with Fig. 8, was explored further with Figs. 19 and 20, which were shown in that order to a single group of Ss.

In Fig. 19, the two halves of the field are equally asymmetrical; neither seems more eligible than the other to continue underneath. Nevertheless, as Table IX shows, their common boundary functioned two-sidedly for only 24% of our Ss; for 71%, it served as the contour of one or the other half, or of each alternately.¹⁰ If our general formulation is valid, this predominant one-sidedness is at least partly due to the irregularity of the field's halves, which, while not completely soluble, is at least mitigated

⁹ When this happens, it should be noted, however, that in the neighborhood of the controversial contour the 'ground' assumes greater surface-quality ('thingness') than it possesses elsewhere.

¹⁰ The fact that such a small proportion of Ss described a fluctuating organization is probably a function of our instructions (see footnote 7).

when one half continues behind the other. This explanation is tested by Fig. 20, each of whose halves constitutes a stable, regular form when seen to end at the common boundary. Here, as Table IX shows, the proportion of Ss for whom the contour is two-sided goes up to 81%. This confirms that 'improvability' of the area continuing underneath is an important factor in one-sidedness; still, for 19% of the Ss even Fig. 20 resolves itself into overlapping planes. Thus, even when the adjacent areas are both

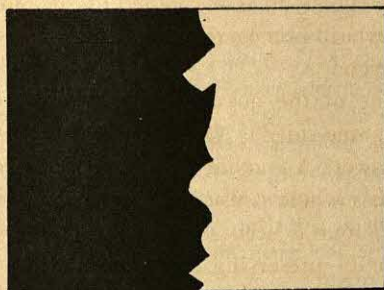


FIG. 19. BLACK AND WHITE AREAS
(More frequently seen as overlapping)

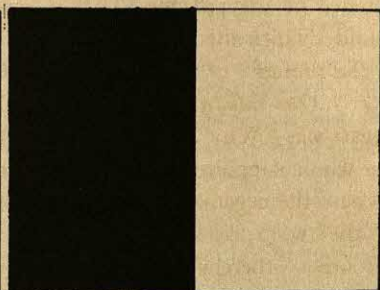


FIG. 20. BLACK AND WHITE AREAS
(More frequently seen as non-overlapping)

TABLE IX

DISTRIBUTION IN PERCENTAGE OF TYPES OF DESCRIPTIONS OF FIGS. 19 AND 20
($N=21$)

Type of description	Figs.	
	19	20
Overlapping planes:		
white area on top	24	14
black area on top	33	5
fluctuating overlapping	14	0
Total	71	19
Contour shared	24	81
Ambiguous	5	0

regular, and equally so, there remains some tendency to see them as overlapping rather than as lying side by side. Our data do not tell us why.

Introspectively, the tension implicit in the shared-contour organization seems to be a matter of less-than-optimal positional arrangement of neighboring forms (see Fig. 3). This is, however, one of many questions related to our problem which can be clarified only by further study. The following section suggests additional questions.

QUESTIONS FOR FURTHER STUDY

(1) *Relations among structural principles.* We have referred above to such familiar principles as closure or completeness, symmetry, simplicity, similarity, unity, but our experiments were not set up to tell us conclusively for each figure just *which* principle is responsible for our results, or what the relative influence of several interacting principles may be. Our method, however, could be adapted to answer such questions.

Consider Fig. 8, for instance; one might vary the structure of the area southeast of the crucial contour to see how much its present potential similarity to its neighbor contributes to the frequency with which the latter is perceived as a square despite the presence of the 'L' shapes. Or one might move the 'L' shapes further apart to see how much their present proximity contributes to the frequency with which the crucial contour is shared—presumably as a function of the laws of similarity and symmetry.

(2) *Nature of the role of outlying areas.* Another set of questions centers around the size of the effective visual field. We have shown that areas not contiguous with the crucial contour can affect phenomenal overlapping. Further experiments could vary the scope of potentially relevant outlying areas; or systematically oppose the requirements of broader visual structures to the requirements of more local ones so as to observe the interaction between them.

In this connection, some of our findings bear on the Ratoosh-Helmholtz hypothesis, which proposes, as the "sole determiner of relative distance,"¹¹ the principle that of two adjacent areas the one whose contour is at the point of intersection more continuous with the crucial contour will appear in front. This hypothesis could be offered, for instance, to explain the responses to our Figs. 2, 4, and 5. Figs. 3 and 8, on the other hand, yield shared-contour responses where Ratoosh would predict one-sided (*i.e.* overlapping) organizations. The differences, moreover, between Figs. 11 and 12, or 15 and 16, cannot be explained in terms of local continuity of contours.

Chapanis and McCleary have also offered a number of instances of unambiguous overlapping where the local 'cues' in Ratoosh's sense were ambiguous, and of shared-contour organizations where local continuity would clearly make for overlapping. They have, furthermore, even given instances of overlapping in a direction exactly *opposite* to what would be predicted on the basis of local continuity. It is interesting, however, that in the latter cases the local detail constituting the overruled continuity was, while clearly visible, small in relation to the size, and irrelevant to the essential shape, of the total area;¹² also that, although on the whole overruled, these small and irrelevant local shapes did substantially subtract from the effect of the pattern as a whole.

The Ratoosh-Helmholtz hypothesis, then, points to a very powerful local factor, involving small portions of the visual field which can exert an influence out of all proportion to their sheer size. Clearly this influence, though impressive, is not the 'sole determiner' of overlapping. Broader structural requirements can make for a decisive outcome where this local factor is indecisive, or for an indecisive outcome where this local factor is decisive; they can even, under unusual conditions,

¹¹ Ratoosh *op. cit.*, 258.

¹² Chapanis and McCleary *op. cit.*, 123-125.

so overrule the local factor that a directly opposite spatial organization occurs. Thus the problem for further research is not to decide between (a), the local influence of contour-continuity at intersection points, and (b), the broader influence of the drawing as a whole, but rather to determine how (a) and (b) interact.

(3) *Factor of familiarity.* It might be pointed out that some of the visual organizations which we have referred to above as symmetrical, simple, consonant with the main directions of visual space, etc., are also more *familiar* to the Ss than alternate organizations would be. Chapanis and McCleary cite familiarity as a separate factor affecting phenomenal overlapping, and Gibson suggests that it may even underly the influence of such factors as completeness and regularity.

Since for good physical reasons symmetrical, simple, etc., forms do tend also to be familiar, special experimental steps must be taken to test the influence of past experience. One question to be answered by such experiments is this: To what extent is familiarity a matter of membership in a temporally extended group of units, corresponding functionally to membership in a spatially extended group (Fig. 8) and reducible to the same laws of organization? What other kinds of processes does the factor of familiarity involve?

SUMMARY

Starting with the questions: Why is a contour usually one-sided? What determines to which side the contour will belong? we constructed a series of visual figures to test the following hypotheses. Phenomenal overlapping is a function of such structural principles as symmetry, simplicity, and similarity. The structure whose organization determines the direction of overlapping includes the area claiming the contour, the area continuing behind, and additional, non-contiguous, areas; the main directions of visual space are also relevant. This formulation was confirmed by the responses of assembled groups of Ss (college students) who wrote descriptions of the figures. Our results also suggest links between the factors determining overlapping, the factors determining three-dimensionality, and the factors determining whether lines will function as figures in themselves or as contours.

We concluded that the figure-ground distinction should be viewed as a special case of phenomenal overlapping: when 'stuff' continues behind the contour—as it usually but not always does—it thereby maintains its homogeneity; when a 'thing' does so, it improves either its own local form or its relation with other parts of the visual field. The general reason for the one-sidedness of contours seems to be that the organization of the field is usually enhanced if one of the two areas bordering on a contour is seen as continuing behind the other. Two-sided (shared) contours do occur when the organization of the field demands them; but they are clearly non-preferred, and the reason for this remains a question for further research. Directions for further research are suggested.

PERCEPTION AS A FUNCTION OF RETINAL LOCUS AND ATTENTION

By WOODBURN HERON, McGill University

The origin of the present study was the series of experiments initiated by Mishkin and Forgays.¹ These investigators found that when English words were tachistoscopically exposed either in the left or right visual fields they were more easily recognized in the right one. They also discovered that when persons capable of reading both English and Jewish were used as subjects (Ss), and when the stimulus-material consisted of words of both languages, Jewish words were more easily recognized when they lay to the left of the fixation-point, and English words when they lay to the right. This occurred under conditions in which S did not know in which visual field the next stimulus-word would appear, or whether it would be an English or a Jewish word.

Two other studies of this problem have been carried out. Orbach showed that Mishkin and Forgays' results with bilingual Ss could be duplicated only if the Ss had learned to read Jewish *before* English.² If English were the prior language, more Jewish words were recognized to the *right* of the fixation-point. Forgays, in a later experiment, used English-reading school children and demonstrated that the difference between scores of left- and right-field was not apparent in Ss who were below Grade VII, while those who were above this grade, however, recognized more words in the right field.³ He suggested that the difference does not appear until S has had considerable experience in reading.

All of these investigators interpreted their results as indicating that since English is read from left to right and Jewish is read in the opposite direction, training establishes a "more efficient neural organization" in the left cerebral hemisphere for English words, and in the right one for Jewish. Their results, they felt, were not in agreement with Lashley's theory of equipotentiality.

These studies raise several points. First, there is some evidence which suggests that English words may not always be recognized more easily in the right visual field than in the left. Glanville and Dallenbach, exposing two rows of letters

* Received for publication February 10, 1956. This study, based on a Ph.D. dissertation submitted to the department of psychology, McGill University, was directed by Professors D. O. Hebb and G. A. Ferguson. The research was supported in part by grants from the Dominion Provincial Mental Health program, and from the National Research Council.

¹ Mortimer Mishkin and D. G. Forgays, Word recognition as a function of retinal locus, *J. exp. Psychol.*, 43, 1952, 43-48.

² Jacob Orbach, Retinal locus as a factor in recognition of visually perceived words, this JOURNAL, 65, 1953, 555-562.

³ D. G. Forgays, The development of a differential word recognition, *J. exp. Psychol.*, 45, 1952, 165-168.

across the visual field, found that letters to the left of the upper row were most accurately perceived.⁴ Crosland has shown that when nonsense-words are exposed in the center of the field, more letters are accurately reported to the left of the fixation-point than to the right,⁵ a finding which seems to be in agreement with the results of similar experiments carried out in German laboratories, which have been summarized by Woodworth.⁶ In addition, Anderson states that under similar conditions, if English and Hebrew nonsense-words are randomly interspersed, bilingual Ss will recognize more English letters to the left of fixation, and more Hebrew letters to the right.⁷ In spite of the fact that letters instead of words were used in these investigations, and a slightly different experimental procedure was followed, it is hard to reconcile these results with the interpretation proposed by Mishkin and Forgays.

Secondly, in reading presumably as many words fall in the right visual field as in the left one, and consequently the right half of each retina must receive as much stimulation as the left. As Mishkin and Forgays realized, if their interpretations were correct any difference in the ease with which words are recognized in the two fields must be due to some attentional process during reading which makes English words in the right field cause more 'learning' than those in the left field. Thus, it is obviously important to find out more about these attentional processes.

A third point concerns the specificity of the visual habit involved; while the difference between the fields exists for words, we do not know whether it exists when non-alphabetical forms are used as stimuli.

The present study consists of five experiments: Experiments I and II deal with the effects of form and of set; and Experiments III and IV with the effects of 'set' on the perception of letters. Experiment V was conducted to throw light on the conflicting results reported by Crosland and Anderson.

METHOD AND PROCEDURE

Subjects. The Ss, all college students, could read English fluently. Three had some knowledge of Jewish, but none was proficient in this language.

Apparatus and procedure. A projection tachistoscope was used, and the exposure-time was approximately 100 m.sec. Letters were typewritten in upper-case elite type, and non-alphabetical material was drawn in India ink on slides. These stimulus-materials were projected on to a white screen, which was uniformly illuminated by two 60-w. bulbs situated 4 ft. away from it. The fixation-point was a dot placed in the middle of the screen. S sat 7 ft. away, a chin-rest being used to control

⁴ A. D. Glanville and K. M. Dallenbach, The range of attention, this JOURNAL, 41, 1929, 207-236.

⁵ H. R. Crosland, Letter-position effects, in the range of attention experiment, as affected by the number of letters in each exposure, *J. exp. Psychol.*, 14, 1931, 477-507.

⁶ R. S. Woodworth, *Experimental Psychology*, 1938, 742 f.

⁷ I. H. Anderson, The effect of letter position on the range of apprehension scores, with special reference to reading disability, *Univ. Mich. School of Educ. Bull.*, 18, 1946, 37-40.

distance and to reduce head-movements. Before the test-series was commenced, S was given 20 exposures as preliminary training, enough to make him familiar with the experimental materials, procedures, and apparatus. He was told to maintain his fixation-point, after he was given a preexposure 'ready' signal, until the exposure actually occurred. When letters were used as stimulus-objects, S wrote those he had seen after each presentation. When non-alphabetical material was used, he was tested by the recognition-method, using a multiple-choice technique.

EXPERIMENTS I AND II

Experiments I and II were conducted to discover whether the superiority of right-field scores was apparent when non-alphabetical material was used.

In both, outlines of simple forms were exposed in random order at angular distances of $3^{\circ} 43'$ to the right and to the left of the fixation-point. Each form appeared as often in the right as in the left field. In Experiment I, in which 15 Ss were used, 44 nonsense forms were exposed in each position; in Experiment II, with 14 Ss, the stimulus-objects were 12 familiar forms, *e.g.* heart, star, triangle, square. The forms subtended a visual angle of $1^{\circ} 30'$.

Results. The results, summarized in Table I, show that there was no

TABLE I
PERCEPTION OF FORMS IN THE RIGHT AND LEFT VISUAL FIELDS

Stimulus-objects	Means		<i>t</i>	<i>p</i>
	left	right		
nonsense forms	23.3	23.7	0.46	>.60
familiar forms	6.0	6.6	1.01	>.30

significant difference between the mean scores for forms exposed in the right field and those exposed in the left.

EXPERIMENTS III AND IV

Experiments III and IV were designed to answer two questions: (1) Is there any difference in scores between the two fields when letters are used as stimulus-materials? (2) What are the effects of controlling set? In other words, if S knows where the letters are to be exposed, will he make a better score than if he does not? Also, are letters as easily recognized in the left as in the right field under these conditions?

In Experiment III, 20 Ss were used. Groups of four upper-case letters arranged in a square were exposed at five different positions to the right, and at five similar positions to the left of the fixation-point along the horizontal meridian. These positions were so situated that the mid-point of the square of letters lay at angular distances of $1^{\circ} 15'$, $2^{\circ} 45'$, $4^{\circ} 15'$, $5^{\circ} 55'$, and $7^{\circ} 7'$ from the fixation-point.

Every letter square subtended $1^{\circ} 27'$. During the course of the experiment, 16 groups of letters were exposed in each field at each angular distance, every letter appearing in the right visual field as often as it did in the left. For half of the exposures, *S* was told on which side of the fixation-point the letters would appear (informed condition); for the other half he was not told (uninformed condition). The letter-groups exposed in the 'informed' condition for one *S* were used in the 'uninformed' condition for the next *S*.

Experiment IV was similarly designed. Twelve *Ss* were used, and the stimulus-objects consisted of single upper-case letters. These were exposed in four different positions along the horizontal meridian in each visual field, at angular distances

TABLE II

RECOGNITION-SCORES FOR LETTER-GROUPS EXPOSED IN THE RIGHT AND LEFT VISUAL FIELDS AT FIVE DIFFERENT VISUAL ANGLES FROM THE FIXATION-POINT UNDER INFORMED AND UNINFORMED CONDITIONS

(*L*=left; *R*=right; *I*=informed; *U*=uninformed)

Distance from fixation-point	Means				Total		t-values*					
	LU	LI	RU	RI	L	R	LI-LU	RI-RU	RI-LI	RU-LU	R-L	
(1) $1^{\circ} 15'$	25.1	25.9	24.8	25.4	51.0	50.2	0.93	0.53	0.87	0.14	0.61	
(2) $2^{\circ} 45'$	16.5	18.2	20.9	20.9	34.7	41.8	1.77	0.08	6.44	6.08	7.08	
(3) $4^{\circ} 15'$	13.3	15.1	15.5	15.4	28.4	30.9	1.79	0.13	0.69	2.11	2.05	
(4) $5^{\circ} 55'$	10.2	10.7	10.7	11.4	20.9	22.1	0.25	1.33	1.51	0.34	1.27	
(5) 7°	5.8	5.5	7.4	6.9	11.3	14.3	0.11	0.10	1.67	0.82	2.36	
(2) plus (3)	29.9	33.2	36.4	36.3	63.1	72.1	2.34	0.10	3.33	5.48	5.68	
All positions	70.9	75.4	79.3	80.0	146.3	159.3	1.45	0.39	3.33	3.15	3.17	

* A *t* of 2.09 is significant at the 5% level; of 2.86, at the 1% level.

TABLE III

RECOGNITION-SCORES FOR SINGLE LETTERS AT DIFFERENT ANGULAR DISTANCES FROM THE FIXATION-POINT, UNDER INFORMED AND UNINFORMED CONDITIONS

Distance from fixation-point	Means				Total		t-values*					
	LU	LI	RU	RI	L	R	LI-LU	RI-RU	RI-LI	RU-LU	R-L	
(1) 4°	12.5	12.7	12.9	12.5	25.2	25.4	0.31	0.58	0.90	0.24	0.24	
(2) 5°	11.5	10.7	12.8	12.5	22.2	25.3	1.79	1.47	1.70	2.81	2.41	
(3) 6°	9.8	9.4	10.7	10.6	19.2	21.3	0.63	0.16	0.16	1.39	1.30	
(4) 7°	7.8	7.4	8.1	8.2	15.2	16.3	0.10	0.26	0.26	1.13	0.07	
(2) plus (3)	21.3	20.1	23.5	23.1	41.4	46.6	1.93	0.93	1.60	2.94	2.35	
All positions	41.6	40.2	44.5	43.8	81.8	88.3	1.17	0.05	1.17	2.04	1.70	

* A *t* of 2.201 is significant at the 5% level.

of 4° , 5° , 6° and 7° from the fixation-point. Sixty-four exposures were given in each position; 32 to the right, and 32 to the left. Half of the exposures occurred under 'informed' and half under 'uninformed' conditions. Every letter appeared as often in the left field as it did in the right one in each position, and under each condition.

Results. Scores were based on the number of letters identified. The results are summarized in Tables II and III, and in Fig. 1. It can be seen that the results of both experiments are similar. More letters were perceived in the right field than in the left, and this difference is most marked

at certain angular distances from the fixation-point. In both tables scores have been pooled for the two positions at which the difference between the two fields is most marked (5° and 6° in the case of single letters, $2^{\circ} 45'$ and $4^{\circ} 15'$ in the case of letter groups). Scores for the left field are significantly greater under the 'informed' than under the 'uninformed'

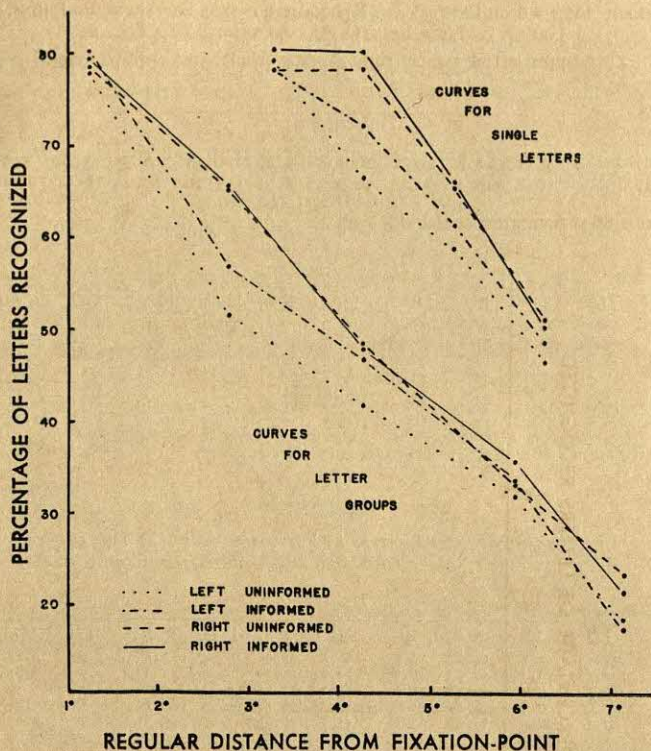


FIG. 1. PERCENTAGE OF LETTERS RECOGNIZED AT DIFFERENT ANGULAR DISTANCES FROM THE FIXATION-POINT.

conditions at certain positions, while those for the right field remain the same whether *S* is informed or not.

One point of considerable interest arose from the results of Experiment III in which groups of letters were exposed. Each letter-group, it will be recalled, consisted of four letters arranged in a square. When reporting what they had seen, the *Ss* tended to write first the letter which appeared in the upper left corner of the square, next the one in the upper right, the one in the lower left after this and finally the letter in the lower right

corner. This occurred no matter whether the group was exposed in the left visual field or the right. Moreover, the Ss reported that this was the way that they saw the letters. That is, it seemed to them that they attended to the letter in the upper left corner of the square first, and the one in the lower right last. This impression was so vivid to some Ss that they said that

TABLE IV

ACCURACY WITH WHICH LETTERS ARE RECOGNIZED IN THE DIFFERENT POSITIONS, WHEN GROUPS OF FOUR LETTERS ARE ARRANGED IN A SQUARE

(1=upper left; 2=upper right; 3=lower left; and 4=lower right)

Field	Means				t-values*					
	1	2	3	4	1-2	1-3	1-4	2-3	2-4	3-4
Left	48.1	41.8	29.8	26.0	3.46	6.35	6.95	3.80	5.39	1.95
Right	50.8	46.2	32.0	30.5	2.70	11.26	10.43	6.60	9.18	0.84

* A t of 2.86 is significant at the 1% level.

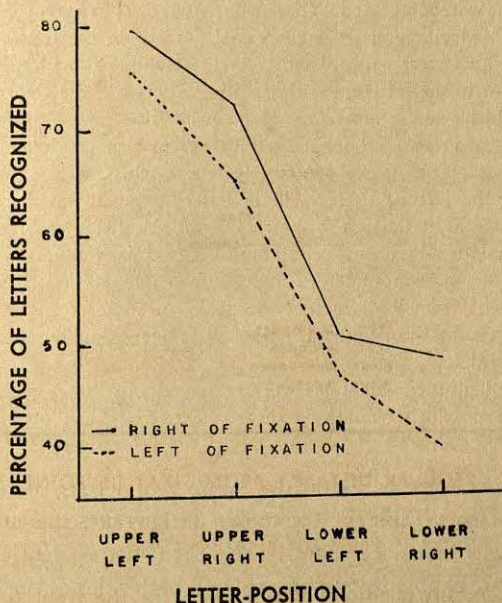


FIG. 2. PERCENTAGE OF LETTERS RECOGNIZED IN EACH POSITION WHEN FOUR LETTERS ARE ARRANGED IN A SQUARE

it felt as if they were fixating each letter in turn—though this, of course, could not occur with the 100 m.sec. exposure.

These reports are borne out by analysis of the data. It can be seen from Fig. 2 and Table IV that the letter which is most often reported correctly

is the one in the upper left corner of the square, while the one in the lower right corner is least often recognized, whether exposure occurs in the right field or the left one.⁸

EXPERIMENT V

The preceding experiments demonstrated that alphabetical material is more easily recognized in the right visual field than the left one when the stimulus appears either in one field or the other. On the other hand, *S* seemed to recognize letters at the beginning of a series more accurately than those at the end, a finding which is in agreement with the results of Crosland and Anderson. In the next experiment, an attempt was made to find out which of these conditions is more important in determining the subject's responses.

Procedure. Fifteen *Ss* were used in this part of the study. Six different types of stimulus-patterns were employed. These consisted of (1) groups of four letters which were exposed either in the right visual field or in the left one; (2) groups of four letters which were exposed either in the right field with a thick horizontal line in the left, or in the left field with the line exposed in the right; (3) two four-letter groups which were exposed simultaneously, one in the right field and the other in the left; (4) two groups as in (3) except that the letters in each group were spaced twice as far apart; (5) single letters which were presented either in the right or in the left field; (6) single letters exposed simultaneously in both fields. These stimulus-materials are shown diagrammatically in Table V. The experiment was so designed that each letter appeared as many times to the right as it did to the left of fixation under each experimental condition. The single letters were exposed at 4° from the center of the field. An angle of $1^\circ 57'$ was subtended by each group, except those in which the letters were double-spaced. In this case the angle subtended was $3^\circ 29'$.

The various stimulus-combinations were presented in random order to ensure that *S* would not know which would be exposed next. Thus, any pre-exposure 'set' on *S*'s part was eliminated.

Results. Scores were based on the number of letters correctly reported. It can be seen from Table V that the conditions of stimulation determine whether the letters will be more clearly seen in the right field or in the left. If letters are exposed simultaneously on both sides of the fixation-point, more are recognized in the *left* field. If, on the other hand, they are exposed in only one field at a time, more letters are seen to the right of

⁸ Recent experiments by Miss Rhoda Cohen have demonstrated that similar results are obtained even if the fixation-point lies in the middle of the square of letters, so that letters on the left fall on the right hemi-retinae, while those on the right fall on the left hemi-retinae. See also Glanville and Dallenbach, The range of attention, this JOURNAL, 41, 1929, 207-236.

fixation. When the letters are double-spaced, the superiority of the scores in the left field is greater than with single-spaced letter-groups, perhaps because the double-spaced letter-groups are not separated from each other as much as the single-spaced ones (see Table V), and thus appear more like a single eight-letter group. In the case in which letters are exposed in one field while a horizontal line is exposed in the other, more letters

TABLE V
COMPARISON OF THE EFFECTS OF SIMULTANEOUS AND SUCCESSIVE PRESENTATION

Type of stimulus*	Means		t	p
	left	right		
(1) xxxx . or . xxxx	13.5	16.6	5.25	<.001
(2) xxx- - - - - or - - - - - xxxx	12.1	15.7	3.63	<.01
(3) xxxx . xxxx	11.5	9.4	1.91	<.10
(4) x x x x . x x x x	14.4	9.7	2.70	<.02
(5) x . or . x	5.5	5.3	0.89	>.30
(6) x . x	5.1	4.6	1.74	>.10

* The dots in the center of this column represent the position of the fixation-point.

were seen in the right field, which indicates that the simultaneous stimulation of the right field must be of a rather specific type to enhance scores in the left field.

In general, the Ss reported that when the letters were exposed in both fields simultaneously, the ones in the left field seemed to 'stand out' more than those to the right of fixation. They reported that it seemed to them they were attending to each of the letters in turn, starting with those at the left.

DISCUSSION

The fact that more letters are recognized in the left field than in the right when English letters are exposed in both fields simultaneously does not seem to be in accord with Mishkin and Forgay's hypothesis that "a more efficient neural organization" is established in the left cerebral hemisphere for English words. Some interpretation is needed which can deal with the results obtained when letters are presented in both fields simultaneously, as well as when they are presented successively. The only type of explanation which seems to account for all the phenomena is one in terms of attentional processes.

If we consider the problem in these terms, the results indicate that we must be dealing with two types of process. The first, which we may conveniently refer to as 'set,' is involved when *S* concentrates on one part of the field, expecting the stimulus-objects to appear there (the 'controlled' conditions of the present study, for example); the second operates *after* the stimulus-material has been exposed, and is affected by the properties of the stimulus. This will be called the 'post-exposure process.'

Considering set first: the experiment in which letters were exposed under 'informed' and 'uninformed' conditions suggests that 'set' operates differentially in the two fields. When *S* knew on which side of the fixation-point the letters would appear, his scores for the right field were no better than when he did not. His scores for the left field, however, were increased significantly. This result suggests that one factor involved in the more efficient recognition of letters in the right visual field is a tendency for the uninstructed *S* to attend to the right.

In most recent tachistoscopic experiments, 'set' is the only type of attentional process which is taken into account. It is generally believed to act before the stimulus occurs, predisposing *S* to perceive or react in a certain way. The idea that attentional processes act after the stimulus has been exposed, however, is a reasonable one. It is obvious that the neural activity involved in perception must persist for some time after the stimulus has been presented. During this period it would be possible for the 'post-exposure' attentional process to operate.

The most noticeable feature of this process, as the *Ss*' report and their objective results indicate, is that the exposed letters are attended to in the order that they would normally be read: letters which would tend to be fixated first under normal reading conditions have their traces 'scanned' first. Thus, there appears to be a close relationship between the eye-movements, or tendencies toward them, established by reading and the post-exposure process.

If tendencies toward eye-movement are important in determining how the post-exposural process operates it is possible to see how the apparently contradictory results obtained under conditions of successive and simultaneous presentation can be reconciled. We know that in reading English there are two main types of eye-movement. The first is a series of short movements from left to right along the line of print, the second consists of movements from right to left at the end of each line. Thus the fluent English-reader presumably has two tendencies established; faced with a line of print there is one tendency to fixate near the beginning of the line and another to move the eyes along it from left to right.

When alphabetical material is exposed in the right field alone, the two tendencies would be acting together. When, however, it is exposed in the left field alone, the tendency to move the eyes to the beginning of the line (presumably the dominant one) would be in conflict with the tendency to move the eyes from left to right. Under conditions of successive presentation we should therefore expect that more letters would be recognized in the right field. When exposure occurs simultaneously in both fields, on the other hand, the dominant tendency to move the eyes to the beginning of the line would result in more letters being recognized in the left field. Familiar and unfamiliar forms would be recognized equally easily in both fields, since, as one does not usually read lines of nonsense figures or geometrical forms there would not be the same tendencies toward eye-movement established for figures as there are for letters. It should be noted, however, that Dallenbach and his co-workers report that when circular stimulus-patches are exposed in both fields simultaneously the patch in the left field is more attentive (clearer or more vivid) for right-handed observers.⁹ The stimuli in these experiments were, however, more peripherally situated than those in the present study.

There is one other point which should be noticed about the way in which the scanning process operates. It will be remembered that when letters are exposed in one field while a solid line was simultaneously exposed in the other more letters were seen to the right than to the left of the fixation-point, indicating that stimulation in the right field must be of a rather specific nature to enhance the scores of the left field. This suggests that before the post-exposural process operates discrimination of the stimulus-materials must occur at some level in the system.

In discussing the scanning which is involved in the selective recognition of letters we appear to be approaching one of the major problems of psychology, the temporal integration of thought and action. As Lashley has pointed out,¹⁰ the phenomena of serial order cannot be accounted for neurologically in terms of associative chains alone, nor in terms of the sensory control of movements. He argues that there must be some process whereby traces and motor neurons can be scanned in sequence. The post-

⁹ K. M. Dallenbach, *Positive vs. intensity as a determinant of clearness*, this JOURNAL, 34, 1923, 282-286; R. S. Burke and K. M. Dallenbach, *Position vs. intensity as a determinant of attention of left-handed observers*, *ibid.*, 35, 1924, 267-269; A. M. White and K. M. Dallenbach, *Position vs. intensity as a determinant of the attention of left-handed observers*, *ibid.*, 44, 1932, 175-179.

¹⁰ K. S. Lashley, *The problem of serial order in behavior*, in L. A. Jeffress (ed.), *Cerebral Mechanism in Behavior: Hixon Symposium*, 1951, 112-135.

exposural process postulated seems to be related to such a mechanism, and the experimental data throw some light on its nature.

It would be premature to construct an elaborate neurologic model to account for the results of this experiment. It is clear, however, that some interpretation along the lines suggested by Hebb could be made.¹¹ The basic assumptions would be: (1) that traces (cell assemblies) corresponding to the various letters are established in the brain, and that these assemblies have motor components, so that the process determining eye movements could facilitate assemblies that were sensory in origin; (2) that these cell assemblies are excited when their corresponding letters are exposed, but that the excitation decays with time; (3) that the excitation produced by the brief stimulation is not great enough to produce a response without additional facilitation,¹² (4) that the excited assemblies facilitate each other sequentially; and (5) that eye movements made during the development of the reading habit are important in determining the sequence.

CONCLUSIONS

This study, then, suggests that the fact that responses may be more easily elicited from one part of the retina than another, which lies in a region of equal acuity, does not necessarily imply that more 'learning' has occurred in some parts of the visual system than in others. The results are not, as previous authors have suggested, incompatible with Lashley's ideas, in view of his preoccupation with the problem of serial order, a problem which is again raised by the present findings. Interpretation of the phenomena seems to be most easily made in terms of an attentional process which is developed during the course of learning to read, and in which eye-movements play a vital role, somewhat along the lines suggested by Hebb.

¹¹ D. O. Hebb, *The Organization of Behavior*, 1949, 60-106.

¹² Assumptions (2) and (3) would be relevant to the observation that Ss frequently state that they 'saw' the exposed stimulus-material clearly, yet are unable to report it. Compare Glanville and Dallenbach, *The range of attention*, this JOURNAL, 41, 1929, 207-236., who state that "clearness of experience cannot be correlated with accuracy of report" (p. 222).

THE RELATIONSHIP BETWEEN A RELATIVELY COMPLEX MOTOR SKILL AND ITS COMPONENTS

By EDWARD A. BILODEAU, Tulane University

The present study attempts to analyze certain relationships between two-hand and one-hand tracking performance. In two-hand, or *total-task*, tracking, the Ss may respond simultaneously with left and right hands; in one-hand, or *part-task*, tracking, the Ss use only one hand. The scores obtained by groups of Ss performing in the part-tasks were used to predict the scores of the total-task of other Ss. The predictions were made by taking advantage of the geometric relationships between the total- and part-tasks.

Analyses of this type are important for understanding the structure or composition of certain complex skills. Over the long run, the general aim is to formulate laws by means of which the composition of relatively complex psychomotor activities can be described in terms of part-skills and relations between part-skills. Though relational terms, or terms of interaction, have been assumed by practically all S-R theorists, no notable data bearing on this issue appear to exist in the field of motor skills.

One way of proceeding would be to combine an S's part-scores in some fashion, and then to compare this predicted score with the score of his total-task. Such a method would, of course, deal with the prediction of individual differences in total performance by means of observations of individual differences in separately practiced, component performances. The present treatment is different, the study by the use of individual differences still being incomplete. The method used here was to combine the part-scores of separate groups of Ss in order to estimate what other groups would do on the total task. Thus, a test is made of how successful is the prediction of two-hand performance from data on the difficulty of separately performed constituent responses when the part-scores are combined in a simple additive fashion. The adequacy of the combining rule is then evaluated in terms of the part-part interactions suggested by the data.

METHOD

Apparatus. The basic apparatus was Model CM101B of the Standard Two-Hand Coördination Test modified in several ways in order to change the task from simple pursuit to tracking.¹ The modified task required S to move a pin a distance of 6 in., from one end of a narrow slot to the opposite end, by appropriate turning of two crank handles. If S did not coördinate his two cranking inputs, or turned in the wrong direction, the pin hit the side of the slot, and further turning of the handles was impossible. The modified apparatus, then, has a slot and pin substituted for a moving target and target follower, respectively. The control-indicator relationship remains the same, but the walls of the slot effectively limit the travel of the pin.

* Received for publication May 13, 1955. This research was carried out under the Air Force Personnel and Training Research Center, Project 7707.

¹A. W. Melton (ed.), Apparatus tests. *AAF Aviation Psychology Program Research Report No. 4*, 1947, 177-206.

A stop was placed at each end of the slot. As the pin left an end-stop, a timer was started, and as the pin hit the opposite end-stop, the timer was stopped. The measure of the response was the time in seconds required to traverse the slot.

The apparatus could be rotated about its center, permitting *E* to fix the slot at any angle desired. By definition, at 0° the pin was moved directly away from *S*'s body. Changing the angle of the slot systematically varied the task. This variation can be described in the following ways: one- or two-hand performance; direction of rotating the handles; and input proportion for the left and right hands.

Only one hand was needed to drive the pin along each of the four angles defining the quadrants; all other angles required two-hand cranking. Angles 0° and 180° required only the right hand, turning counterclockwise and clockwise, respectively. Angles 90° and 270° required only the left hand, using clockwise and counterclockwise cranking, respectively. For angles (within the four quadrants) requiring the use of two hands, the direction of handle rotation is implied by that required along the nearer coördinates. For example, at 45° the right hand turned counterclockwise and the left hand turned clockwise. Further, because 45° represents a diagonal, it is also implied that the hands were required to move at equal rates, *i.e.* input proportions of 50% for each hand. For each two-hand angle the 6-in. path of the pin represented the hypotenuse of a right triangle. Solutions for the legs represented magnitudes of input required of the right and left hands. These magnitudes were expressed, however, as the ratio of a leg to the sum of the two legs.

Twenty-four angles, multiples of 15° , were used, making seven possible relative magnitudes of input for each hand, 0, 21, 37, 50, 63, 79, and 100%. These proportions repeated in each quadrant. If the proportion for the right hand was 63%, for example, the right hand had to turn faster than the left, but the left still contributed 37% of the total motion. Whenever a hand contributed more than 50%, that hand will be said to be *predominant*.

Procedure. The *Ss* were 480 basic airmen trainees at Lackland Air Force Base. The *Ss* were assigned to groups of 40, and without known bias, to one of 12 slot angles. Every *S* was required to move the pin through an assigned angle (x°), and on the next trial, to return the pin to the starting position through the opposite angle ($x + 180^\circ$). In other words, the scores for opposite angles are related, but the scores for adjacent angles are unrelated. This means there were 12 groups of *Ss* and 24 treatments.

While instructing *S* about the required directions of rotation, *E* moved the pin through the slot in a demonstration. The *S* was asked to drive the pin from one end to the other *as fast as possible*. Four trials (two for each opposite angle) with 30-sec. rest between trials were administered. One half of each subject-group began with one angle, while the other half began with the opposite angle. The order of administration for treatment groups was counter-balanced and only right-handed *Ss* were tested.

RESULTS

Rank order of one-hand angles. The mean time required by each group having a one-hand task is shown at the top of Table I. These data were used to predict the results from the total tasks. The rank order of the four-part tasks gave the following relationship between hand (*R* and *L*) and direction of rotation, clockwise

and counterclockwise (*cl* and *cc*): $R_{cl} > R_{cc} > L_{cc} > L_{cl}$, where the *sign* ($>$) should be read *superior to*.

This relationship, obtained when inputs of the hands are one and zero, does not necessarily hold for other input proportions. Consequently, for example R_{cl} and L_{cc} may not be the most favorable combination in a two-hand task, nor L_{cl} and R_{cc} the least. To evaluate the worth of a simple combining rule, it was decided, however, to make predictions about the two-hand tasks by weighting the one-hand data by the values of input proportion appropriate to the two-hand angle.

(1) *Predicting the rank order of two-hand angles.* The predicted ranks of

TABLE I
TIME-SCORES, RANKS PREDICTED AND OBTAINED FOR THE TRAVERSAL OF ALL ANGLES
(Time in seconds)
One-hand performance

Angle	Time	Angle	Time
0°	5.79	180°	5.18
90°	8.37	270°	7.79

Two-hand performance

Angle	Rank		Time	Angle	Rank		Time
	predicted	obtained			predicted	obtained	
195°	1	10	8.80	315°	11	5	7.77
165°	2	4	7.48	240°	12	13	10.77
210°	3	1	6.73	300°	13	17	11.40
345°	4	8	8.39	45°	14	6	7.83
150°	5	3	7.29	120°	15	19	12.15
15°	6	16	11.22	255°	16	14	10.91
225°	7	2	7.02	285°	17	18	11.59
330°	8	11	8.82	60°	18	15	10.93
30°	9	7	7.95	105°	19	20	13.48
135°	10	9	8.54	75°	20	12	10.38

two-hand angles were derived from the sums of the weighted time-scores for one-hand angles. A general expression of the procedure is given by the following equation:

$$t_{R, L} = at_R + (1 - a) t_L \dots \dots \dots [1]$$

where t_R and t_L are the times for traversal along the adjacent one-hand coordinates, and a and $1 - a$ represent the input proportions of hands R and L , respectively, for any two-hand angle in question. The angle 195° ranked first, for example, because the largest input proportion (79%) was multiplied by the smallest time score (R_{cl}), and in turn added to the product of the smallest proportion (21%) and the better time for the other hand (L_{cc}).

This simple method of combining part scores involved no rules of interaction and did not place a special emphasis upon those two-hand angles where the inputs were equal. It did, of course, favor all angles in which R was predominant.

This method of combination gave the angles the ranks shown in the lower half of Table I. A value of 64% for ρ was obtained between the predicted ranks and ranks based on the time scores. To understand the correlation better, it will be necessary to examine a few functional relationships.

The mean time in seconds required to traverse the slot has been plotted against the angle in Fig. 1. The connected points represent the data for those angles, separated by 15° , requiring the use of two hands. The x's represent the data for the four part-tasks. The points in the results with two hands are somewhat irregular in trend, but a sine-like ordering over angles can be distinguished. That is, there appear to be maxima in the neighborhood of 90° and 270° and minima in the vicinity of 180° and 360° .

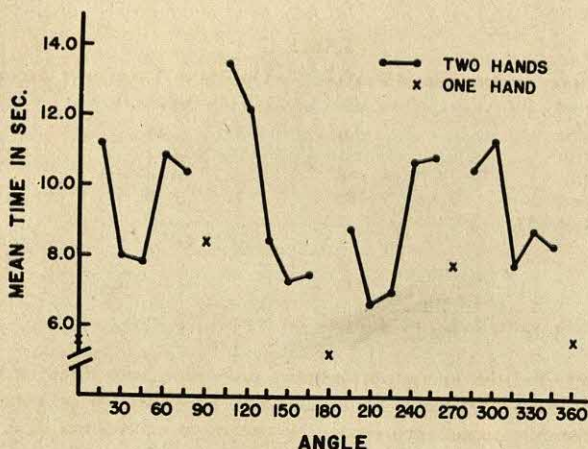


FIG. 1. MEAN TRAVEL-TIME AS A FUNCTION OF ANGLE THROUGH WHICH A PIN IS MOVED BY COÖRDINATED TURNING OF TWO HANDLES

Fig. 1 confirms an expectation from Equation [1] in that some two-hand groups perform better than some one-hand groups. In addition, the longest times of traversal with two hands tend to be associated with angles near the most difficult one-hand angle (90°), and the shortest times of traversal tend to be associated with the easiest one-hand angle (180°), save for two outstanding exceptions. The means for angles 15° and 195° are in opposition to the predicted trend. Both angles require a small input by the left hand. Other minor departures from a sine-like function in Fig. 1 likewise suggest that when the required input reaches the maximum (79%) for one hand and minimum (21%) for the other, travel-time is increased over that anticipated.

(2) *Analysis of one-hand treatments.* The scores of the four one-hand treatments were analyzed according to Lindquist's Type I analysis of variance, summarized in Table II.² This design is appropriate since each *S* served under two conditions of rotation. The main effect of direction of rotation (*cl* vs. *cc*) is not significant, but the main effect of hand (*R* vs. *L*) is statistically significant as is the interaction of the two main effects. The nature of the interaction accounts for the lack of

² E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*. 1953, 267-273.

difference in the main effect of rotation, *i.e.*, *cc* is superior for *L*, but *cl* is superior for *R*. For either hand, direction of rotation is important, for both *ts* are significant far beyond the 1% level of confidence.

(3) *Analysis of two-hand treatments.* The statistical analysis of the 20 two-hand treatments was more involved because of the additional variable of input proportion. The three variables were analyzed according to Lindquist's Type IV analysis of variance.³ Three important components of variance were suggested: input proportion (1% level of confidence), direction of rotation for the left hand (5% level

TABLE II

SUMMARY OF ANALYSIS OF VARIANCE OF ONE-HAND TREATMENT GROUPS			
Source*	df.	MS	F
Between Ss:	79	5.37	
hand (B)	1	269.21	135.63†
error (b)	78	1.98	
Within Ss:	80	38	
Direction (A)	1	.01	.05
AB	1	14.17	66.92†
error (w)	78	.21	
Total	159		

* The notation is after Lindquist. † Significant beyond 1% level.

of confidence), and the interaction of input proportion with direction of rotation for the right hand (1% level of confidence). The direction of rotation for *R* was not statistically significant overall, the interaction of rotation of *R* and input proportion suggesting that *cc* rotation is superior at the smaller values and *cl* rotation at the larger values of input-proportion.

The overall trend for input-proportion showed a decreasing time for values ranging from 21 to 63%, and then an increased time for the larger proportion, a finding previously suggested by inspection of Fig. 1. The same trend held for either hand when considered by itself. For the left hand, however, the initial downward trend was not as pronounced as for the right and, beginning with an input of 50%, the upward trend was more pronounced than with the right hand. Thus, increasing the predominance of *L* or *R* increases response-time, but the effect for *R* is not as great as for *L*. The input-proportion trend is only generally in line with that expected from the simple combinational formula. When the hands contribute equally, or with inputs of 50%, performance is relatively good, though not at its best. The best performance is obtained with *R* rotating clockwise with 63% of the input in combination with *L* rotating counterclockwise with 37% of the input.

DISCUSSION

Results have been presented on speed of one and two-hand behavior, analyzed by (a) hand, (b) direction of rotation, and (c) proportion of the total input required of each hand. A plot of mean travel-time against angles suggested a sinusoidal function, not unlike one reported by Brogden for precision of movement

³ *Ibid.*, 285-288.

at following a target with *one* hand at various angles from the body.⁴ A comparison of the maxima and minima of Brogden's data with the present data indicates systematic differences. The results on speed are, however, less orderly and the relationship is presented only as suggestive. Additional procedural and analytic techniques ought to be instituted before attempting as precise a mathematical description as the one Brogden presents. For example, the present results may be dependent in part upon the plane in which the rotation is accomplished.

The results showed that the simple combinational formula was only fairly successful in predicting the ranks of combinations of one-hand tasks. The statistical analysis of data from relatively unskilled Ss suggests that the performance of one hand does interact with that of the other, a fact not deducible from observations of part skills. Assuming that there are few or no interactions among components of a task for novices is, of course, a relatively naïve view of the composition of total tasks. Nevertheless, it was thought that some estimate of the success of such a method of combination should be obtained. Despite the significant interactions which the data provided, it can be said that, in the absence of information regarding interactions, a simple additive formula yields a useful first approximation.

It is necessary to make clear the manner in which the Ss were unskilled, and what the task would demand of a skilled performer, in order to place the present results along the skill-continuum. All of the Ss used were unpracticed and fell far short of the speed of skilled performers. On the basis of what some unskilled Ss did with one hand, prediction of what other unskilled Ss would do with various combinations of two hands was attempted. When *S* performs on a two-hand angle he makes his forward progress largely by maintaining the input proportions demanded by the angle used. Only very slight progress can be made by a rapid alternation between the hands—a method giving the tracking pin a stepwise course similar to the track of the novice in mirror drawing. Though most novices do alternate, inadvertently or otherwise, it is most unrewarding, for the crank handles jam as the pin binds upon the sides of the narrow slot; blocking with one or two hands is, in turn, likely to lead to setting in the wrong direction of rotation once motion is again commenced. Very quickly, therefore, *S* learns to devote his effort to coördinating his hands to the desired input ratio with a slow rate of cranking; the rate of motion is stepped up later, as *S* gains skill in maintaining the correct input ratio.

To say that the skilled *S* is faster than the unskilled is trivial. It is necessary to know why the time is shorter. It is suggested that the terms of the equation combining part-scores will provide the answer to this question. Research in progress with *well-practiced* Ss suggests that the time taken with two hands is equal to the larger term of Equation [1]. In other words, there appears to be no interaction or interference between the hands, since the prediction can be based upon the data of but a single hand. Such Ss are said to be coördinated because the time of the hand taking the shorter time is completely distributed over the time of the hand taking the longer time. The novice, however, shows much less time-sharing between the hands in that from time to time one hand may be working and the other motion-

⁴ W. J. Brogden, The trigonometric relationship of precision and angle of linear pursuit-movements as a function of amount of practice, this JOURNAL, 66, 1953, 45-56.

less. Even when the hands are working in the proper ratio, say one to one (the diagonals of the present task), this is accomplished by sacrificing speed, the speed which can be demonstrated by presenting but one of the parts. The skilled *S* does not sacrifice speed to maintain the proper ratio of input. Because the behavior of the novice is so varied, no simple combination of part-scores will fully describe his total-task performance, whereas the opposite may be true of the highly skilled.

The predictive method used in the present study does not deal directly with individual differences as such—though that study is necessary. It will be recalled that no *S* who took a two-hand angle also took a one-hand angle. The data required for predicting individual performance in total-task from part-task have been gathered with a small number of exceedingly well-practiced *Ss* and a larger number of semi-skilled *Ss*, and the analysis of them is now underway.

SUMMARY

Unskilled *Ss* performed on tasks requiring the use of two hands (total-tasks), other *Ss* performed on related tasks with one hand alone (part-tasks). The relationship between the scores on the two-hand tasks and those on the one-hand tasks was investigated by taking advantage of certain geometrical relationships characterizing the inputs required of the left and right hand in the two-hand task.

The data suggested that there is a strong relationship between performance with both hands and the performances of the separate hands. The relationship, however, is more than a simple combination of weighted one-hand scores, though such a method can account for some of the variance in two-hand tasks.

The data also suggested the nature of the terms of interaction which might be used to predict total scores more adequately. It was proposed that the problem of total-task prediction will prove to be easier for skilled *Ss* than for the unskilled.

RHYTHMIC DISTURBANCES IN A SIMPLE VISUAL-MOTOR TASK

By E. T. KLEMMER, Operational Applications Laboratory

Several studies of reaction-time have revealed difficulties which occur when stimuli are presented at rates of one each second or oftener.¹ In the words of Woodrow: "I was unable to obtain reactions at a rate faster than one every two seconds, as with very fast rates it becomes very difficult to keep from beating time instead of reacting."² Klemmer was able to get consistent simple reaction-times (*RT*'s) with one stimulus per second after practice, but none of his *S*s consistently could react to two stimuli per second.³ The difficulty usually is apparent to *S* and appears either as a tendency for the responses to creep up on the stimulus and soon precede it, or lag farther and farther behind until an entire stimulus-cycle elapses before a response occurs. During the run, *S* loses track of the relation between individual stimuli and responses and therefore does not know if his present response is a late response to the previous stimulus or an early response to the next stimulus.

In view of the fact that simple *RT*'s seldom exceed 0.25 sec., difficulty with stimuli spaced two to four times that far cannot be easily explained in terms of sequential stimulus-response overlap. Indeed, if the *RT* is lengthened by using a disjunctive (choice) test instead of the one-light-one-key test, the rhythmic difficulty disappears. Thus, we have the paradoxical finding that *S* can do better on the more difficult test requiring discriminative response than on the simple *RT*-test. The present experiment will demonstrate the above findings for simple and disjunctive tasks and attempt to discover whether the rhythmic difficulty is due to stimulus-properties, response-properties, or both. Since the effects are not limited to situations in which speed of reaction is stressed, the approach here is in terms of a more general visual-motor task rather than *RT*-tests. Rhyth-

* Received for publication January 20, 1956. This is AFCRC Technical Note 55-3 of the Operational Applications Laboratory, Air Force Cambridge Research Center, Bolling Air Force Base, Washington 25, D.C.

¹ Herbert Woodrow, The measurement of attention, *Psychol. Monogr.*, 17, 1914, (No. 76) 30-38; C. W. Telford, The refractory phase of voluntary and associative responses, *J. exp. Psychol.*, 14, 1931, 28-29; E. T. Klemmer, Time uncertainty and simple reaction time, *ibid.*, 51, 1956, 179-184.

² Woodrow, *op. cit.*, 31.

³ Klemmer, *op. cit.*, 179.

mic difficulty will be indicated by an inconsistent phase-relation between stimulus and response.

METHOD

Subjects. Five graduate students (men) from Catholic University served as paid Ss.

Apparatus. Each S was seated with his right (preferred) hand resting on a table and the fingers on a set of five telegraph-keys. Immediately in front of the keys was a set of five small neon bulbs arranged in an arc. The stimuli were programmed by teletype-tape, and the time between the successive stimuli and responses was encoded into a binary number and punched on the tape. Responses also were punched on teletype-tape for later scoring. The time-interval of acceptance for any stimulus was not the interval between successive stimuli but rather an interval one cycle (0.52 sec.) long situated around the concentration of responses for that test. This procedure gives S maximal credit for correct responses. In all tests each stimulus-light remained on for 50% of the cycle or 0.26 sec.

Training. All Ss were trained on a five-light-five-key test with single light-flashes presented randomly at the rate of approximately 2 per sec. The Ss were instructed to press the key corresponding to the lighted lamps and to keep up with the lights, since very late responses would not be counted correct. Each S was given repeated runs of 250 stimuli on this training-test until he had achieved a run of 98% correct responses or better.

Experimental tests. The experimental tests were designed to determine what factor in the five-light-five-finger test eliminates the rhythmic or phase-keeping difficulties of the 2 per sec. one-light-one-key test. Three classes of stimuli were used and three classes of response. The simplest stimulus was a single light which flashed regularly every 0.52 sec. The second stimulus-series involved all five lamps flashing one at a time in regular left-right, right-left order: 1 2 3 4 5 4 3 2 1 2 3 and so forth. The third stimulus-condition was the random order of the five lamps.

The three classes of response correspond closely to the stimulus-classes. The simplest response involved a single key pressed once for each light, the second response-class involved five keys pressed one at a time in regular left-right, right-left order, and the third involved five keys pressed one at a time in random order. Note that it is possible to pair non-corresponding stimulus- and response-classes. For example, S may have only one key which he presses for each light even though the lighted bulb is changing position at random.

The six stimulus-response combinations actually used are shown in Table I.⁴ These include the five-light-five-key disjunctive test (Test 1) which required a discriminative response and had already been found to give rise to no rhythmic difficulty, and the one-light-one-key test (Test 6) which gave phase-trouble most of the time. The other four tests aim at discovering whether the phase-difficulty in the simple situation can be overcome by having the stimulus change position regularly (Test 4); or randomly (Test 2); the response change regularly (Test 5); or the stimulus and response both change regularly (Test 3). Note that although stimuli and responses change in these four tests, in no case is S transmitting information

⁴Of the nine possible stimulus-response pairings, six seemed to offer reasonable tasks for S. The other three would compel him to generate a random series of responses or to make incompatible responses.

about the stimulus since he always knows what the stimulus and response will be in advance. Of course, no information is transmitted in the one-light-one-key test either. Each of the five Ss took one run on each test in the order listed in Table I and then one run on each test in inverted order.

The Ss were informed of the stimulus-response correspondence required in each test and were instructed simply to press the proper key for each light after each light. They were cautioned not to get behind, *i.e.* not to let two lights go by before they responded to the first one, but no additional urging toward speed of response was given. It had been found that any emphasis on speed of response would only lead to more errors in Test 1 and more anticipatory responses in the other tests.

RESULTS

Rhythmic or phase-keeping difficulty in these tests affects the time of response relative to the stimulus and thus will appear as a modification of

TABLE I

DESCRIPTION OF THE SIX TESTS EMPLOYED

Each test consisted of two runs of 250 stimuli each. Stimuli were presented at a constant rate of one every 0.52 seconds.

Test	No. of lamps	Order of lamps	No. of keys	Order of keys
1	5	Random	5	Random*
2	5	Random	1	—
3	5	Regular	5	Regular*
4	5	Regular	1	—
5	1	—	5	Regular
6	1	—	1	—

* Ss instructed to follow light-position.

the frequency-distribution of time of response. If *S* maintains a consistent phase-relation with the stimulus throughout the test, his responses will form a peaked distribution relative to the stimulus much as might occur in the usual *RT*-test. If a rhythmic disturbance is present, causing responses to drift or jump in phase-relation with the stimulus, an irregular response-time distribution will result. Under conditions of extreme phase-drift, this distribution will approach a rectangular shape over the entire inter-stimulus interval.

The results of all tests are presented as frequency-distributions of response-times, one of which is illustrated in Fig. 1. The time-axis is quantized in 0.04-sec. units for the frequency-count. Raw frequencies are plotted on the ordinate. Each distribution contains the responses to 250 stimuli—usually a few less than 250 in number (due to failures to respond in some scoring intervals).

Response-distributions for all 60 runs are presented separately in Figs. 2 and 3. Since most of these distributions are far from normal or even

regular, no attempt is made to give means or *SDs*. Evidence for the reliability of the distributions is present in the similarity between the repeated runs of the same *S* on the same test denoted by 'a' and 'b' in Figs. 2 and 3. For most of these 'a'-'b' pairs, the 'b'-run was given on a separate day from the 'a'-run, and several other tests were interpolated.

In Test 1 (Fig. 2) three of the five *Ss* show fairly usual distributions, such as might arise from a typical *RT*-test. Test 1 consisted of following

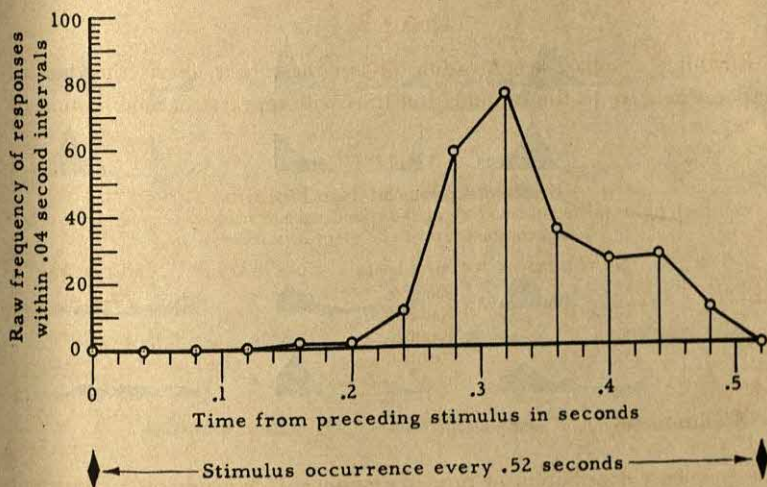


FIG. 1. ILLUSTRATIONS OF THE RESPONSE-TIME DISTRIBUTIONS
PLOTTED IN FIGS. 2 AND 3

The frequency of responses within ± 0.02 sec. of the stimulus appears at both ends of each plot. Each distribution contains approximately 250 responses. The distribution illustrated here is for *S* 1, Test 1, first run.

randomly occurring lights. *Ss* 3 and 5 show fairly concentrated but bimodal distribution in both 'a'- and 'b'-runs. An analysis of the order of response-times for these two *Ss* shows that rapid shifting between the two modes occurred throughout the tests (rather than a single shift in mean toward the middle of the run). Scores of percentage correct are given alongside the distributions. Test 2 involved the same stimuli as Test 1, but the *Ss* had only one response-key. Fig. 2 shows that this simplification of the response led either to irregular distributions over the entire cycle or to a piling up of responses just before the stimulus.

In Test 3, one out of five stimuli occurred in regular order and *S* followed by pressing corresponding keys. Column 3 of Fig. 2 shows that the

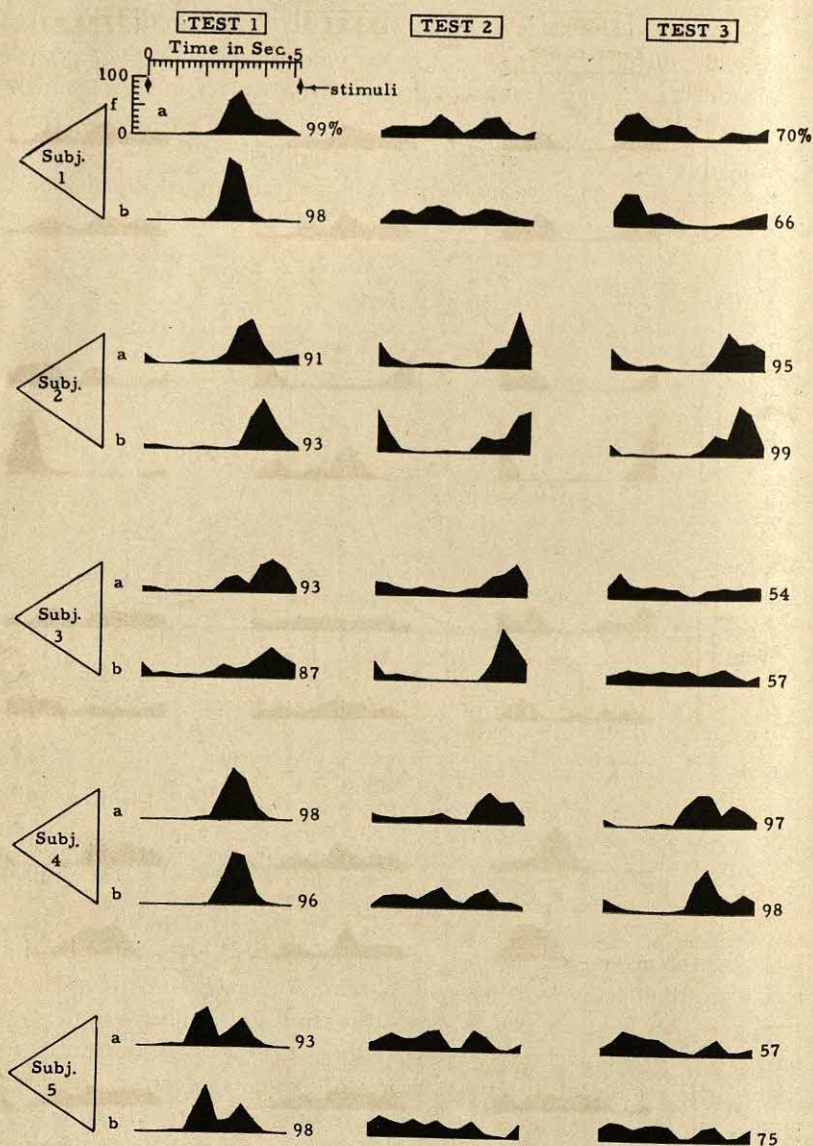


FIG. 2. RESPONSE-TIME DISTRIBUTIONS FOR TESTS 1-3, WITH TWO RUNS FOR EACH OF FIVE Ss
For the distributions of Tests 1 and 3, percentage-correct scores are shown.

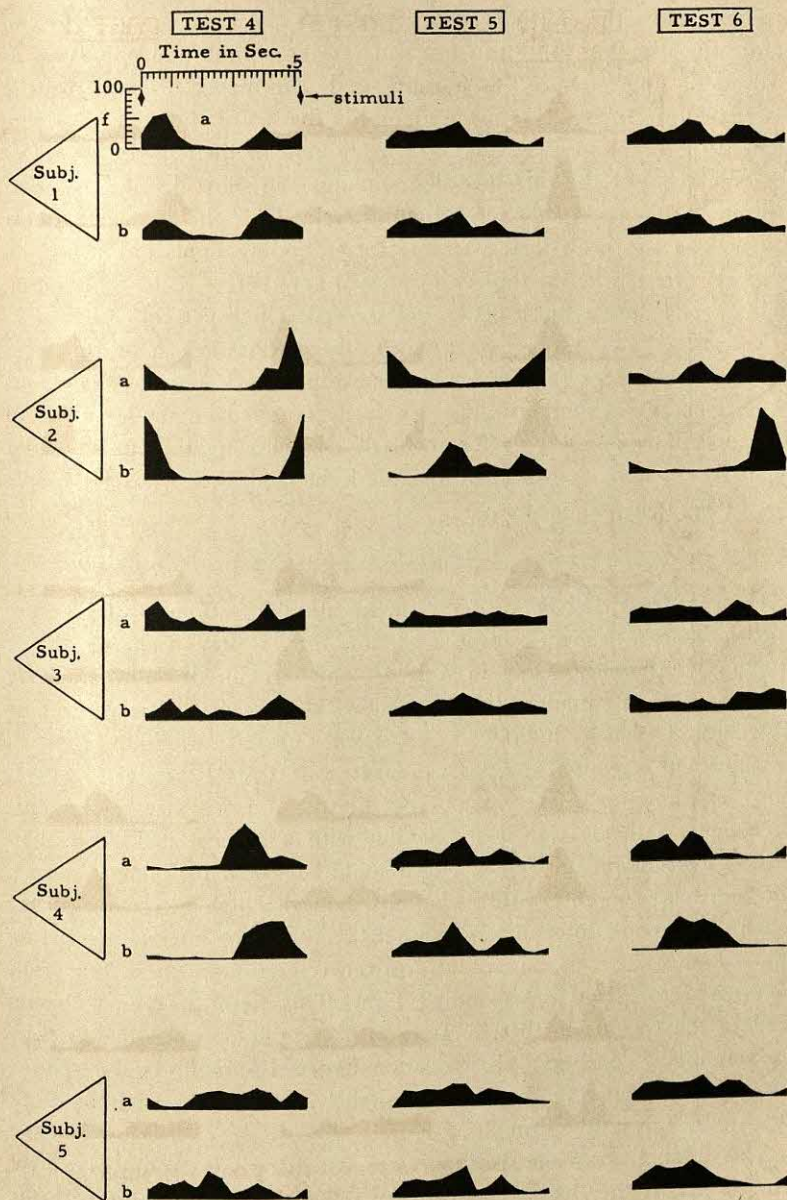


FIG. 3. RESPONSE-TIME DISTRIBUTION FOR TESTS 4-6, WITH TWO RUNS FOR EACH OF FIVE Ss

responses are either fairly well scattered over the cycle or are piled up before the stimulus. Since this test had varying stimuli and responses, an analysis of number-correct is possible, and percentages for an optimally placed scoring interval are given alongside the distributions. The *Ss* who showed the high score (*Ss* 2 and 4) made most of their responses *before* the occurrence of the corresponding stimulus and show distributions concentrated at the end of the inter-stimulus cycle. Thus, the *Ss* able to follow instructions and press keys after the corresponding lights did poorly on number correct, whereas those who pressed keys before the corresponding lights kept in phase much better and showed a high percent-correct.

Tests 4, 5, and 6 (Fig. 3) produce distributions which do not seem different from those for Tests 2 and 3 except that *S* 2 showed two runs with response-distributions almost perfectly centered on the stimulus. In Test 4, *S* hit one key only for regularly changing lights; Test 5 used a regularly changing response for a single light; and Test 6 was the simple case of one light and one key.

DISCUSSION

It is apparent that none of the modifications of stimulus or response tried in Tests 2 through 5 made it possible for *S* to stay in phase with the stimulus consistently. Only in Test 1, which required a discriminative response to each stimulus, did all *Ss* show the concentrated response-time distributions which indicate a consistent phase-relationship between stimulus and response. The phase-difficulties of Tests 2 through 6 do not, however, show up in every run. From Figs. 2 and 3 it may be seen that in every test there is at least one run with a concentrated (consistent phase) distribution. In Test 3, which could be scored for number-correct, the four most concentrated distributions showed a high percentage-correct when the response-interval was so set that most of the interval was before the stimulus. That is, almost all correct responses in these four runs occurred before the corresponding light. This result in Test 2 would lead to the conclusion that in Tests 2, 4, 5, and 6, when the *Ss* were maintaining a consistent phase-relation between stimulus and response, they were anticipating the stimulus slightly, or (in two trials of *S* 2) synchronizing with the stimulus.

This finding raises the question as to whether the *Ss* could have kept phase better if they had been instructed to synchronize response with the stimulus rather than to respond after the stimulus. To answer this question, four *Ss* were given 750 trials each on the one-light-one-key test with

instructions to respond simultaneously with the stimulus. Three of the four developed nearly rectangular distributions and one turned out a concentrated distribution peaked just before stimulus onset. Since these results are not different from Test 6 of Fig. 2, it appears that the phase-keeping difficulties are not due to the particular instructions given. The fact that one *S* could keep phase in the synchronous tests, together with the several peaked distributions in Tests 2-6, shows that the phase-difficulty can be overcome by some *Ss* some of the time at the 2 per-sec. rate in the test not requiring a discriminative response. Only in the test requiring a discriminative position-response, however, were all *Ss* able to stay in phase with the stimulus.

SUMMARY

A series of tests was designed to investigate rhythmic difficulty in key-pressing to flashing lights. All tests were made with a stimulus-repetition rate of 2 per-sec. The stimulus was a single lighted lamp mounted alone, or, in other tests, in a bank of five placed in front of five keys. When five lamps were used, the position of the lighted one changed either randomly or regularly. *S* was instructed to hit one key after each light. In some tests he pressed one out of five keys, in other tests he pressed the same key each time regardless of light position. In only one of the six tests so formed was a discriminative response required, that is, in only one test was *S* required to transmit information about light-position. In all other tests *S* knew in advance what each stimulus and response would be.

The results for five *Ss* on the six tests showed that: (1) Only when there was uncertainty both about the position of the next light and the response to it were all *Ss* able to maintain a consistent phase-relation between stimulus and response. (2) The lack of discriminative response led either to an irregular distribution of responses over the entire inter-stimulus interval or to a piling up of responses close to the stimulus. (3) Instructions to respond simultaneously with the stimulus produced distributions very similar to those for instructions to respond after each stimulus.

THE RELATIVE ACCURACY OF VISUAL PERCEPTION OF MOTION DURING FIXATION AND PURSUIT

By JAMES J. GIBSON, OLIN W. SMITH, ALFRED STEINSCHNEIDER,
and CHARLES W. JOHNSON, Cornell University

The translatory motion of an object may be perceived under two quite different conditions of eye-behavior. When the eyes are fixated on some part of the stationary environment the image of the object moves across the retina of each eye. When the eyes are fixated on the object itself, however, they undergo what is called pursuit-movement and the image of the object does not move across the retina but remains approximately at the fovea. The image of the environment moves across the retina instead.

These two modes of observation are quite different. The former is passive while the latter is active, that is, involves overt motor activity concomitant with the perception. The former requires visual attention to some part of the stationary field while the latter permits visual attention to the object itself. The kinds of retinal stimulation which accompany these modes of observation are likewise different; the former involves motion of a figure within the total retinal image while the latter involves motion of the background image instead of the figure. Introspectively, the impressions of motion obtained are different in the two cases. Nevertheless, the perceptions are in some sense equivalent. The naïve observer sees in both cases an object in motion in a certain direction relative to its environment. How does this equivalence for perception arise? In what respects are these modes of observation not equivalent and in what respects are they equivalent?

An instance of non-equivalence is the report that the apparent speed of a moving object is greater when a stationary point is fixated than when the object itself is fixated. Fleischl made such a report in 1882, concluding in fact that observation with stationary eyes produced an impression of velocity about twice as fast as observation with pursuing eyes.¹ Aubert confirmed the phenomenon and accordingly it has been called the Aubert-Fleischl paradox.² Brown surveyed the literature on the preception of velocity in 1931 and, in one of a series of experiments,

* Received for publication October 14, 1955. Supported in part by the U. S. Air Force under Contract AF18(600)-443.

¹E. Fleischl, *Physiologische-optische Notizen, Sitzsber. Akad. Wiss. Wien*, 86, 1882, 17-25.

²H. Aubert, *Die Bewegungsempfindung, Arch. f. d. ge. Physiol.*, 39, 1886, 347-370, 40, 1887, 459-480.

further confirmed the illusion.³ The method required *O* to compare two moving belts, each presented behind a frame or window, one of which had to be adjusted to the speed of the other. One belt was always observed by fixating the edge of the window and the other by following the pattern on the belt itself. The objective speed of the belt seen with fixation was found to be less than that seen with pursuit when the two were judged equal; in other words, a given speed seen with fixation is phenomenally faster than the same speed seen with pursuit.

The Aubert-Fleischl phenomenon is an error in velocity-perception obtained when the standard and variable are observed in different ways. This procedure, however, is an unnatural one to adopt. What no investigator seems to have done is to determine whether these ways of observing are equally *accurate* for velocity perception. Can *O* match the variable to the standard as well when he fixates both as when he pursues both? This question is important since it concerns the discriminative efficiency of the two modes of ocular behavior.

The two modes of observation will be designated as 'fixation' and 'pursuit' of the eyes. It should be remembered, however, that they have much in common; both serve the function of attentive eye-movements to the objects of a complex environment, whether stationary or moving.

Apparatus and method. The apparatus consisted of two boxes 21 in. square and 7.5 in. deep. The square front of each box was covered with gray cardboard, in the center of which was a glass window, 7 in. high by 9.5 in. wide. Inside each box was an endless belt of wallpaper, stretched over a pair of rollers driven by a variable speed motor. The flat surface of each belt filled its window with a continuous pattern of dense brown and white spots of 1/4 in. diameter. Each box was placed on a table 4 ft. from *O*. Shadowless illumination of the belts came from the two lamps mounted just over *O*'s head. The sight lines from *O* to the center of each window were 70° apart, hence when one was looked at, the other was visible only in extreme peripheral vision. Motion of each belt was always from top to bottom of the window. The speed of each motor was controlled by a Variac mounted in a separate box with its control knob protruding. One such knob was operated by *E* and the other by *O*, but *O*'s knob was hidden from his view so as to prevent knowledge of the setting.

The method of reproduction was used. Each *O* matched the speed of the variable to that of the standard, looking back and forth between them. He was never informed of his error. The speed of the standard was 20 ft. per min., or 4 in. per sec. At a distance of 4 ft., this represents an angular velocity of approximately 4.9° per sec. On half the trials *O* 'fixated,' i.e. fastened his gaze on a small black square glued to the center of each window, and on the other half he 'pursued,' i.e. followed one of the spots on the moving surface with his eyes until it disappeared. On split halves of these trials the variable was either on *O*'s right or on his left. Each of the possible 24 orders was used, one being randomly selected for each *O*.

³ J. F. Brown, The visual perception of velocity, *Psychol. Forsch.*, 14, 1931, 199-323.

Procedure. The instructions given each *O* after being seated were as follows:

This experiment is concerned with the perception of movement. I will set the speed of *this* surface (*E* points) and your task is to adjust the speed of *this* surface until they are equal. Turn the knob to right or left to make it go faster or slower. Alternate between faster and slower until you are satisfied. Any questions?

The instructions for fixation were as follows:

In equating the speeds it is extremely important that you fix your eyes on the little black square and on nothing else. Do not let your eyes follow the moving surface at any time; keep your eyes on the square as if your life depended on it. While fixating it you will get an impression of the speed of the moving surface. Keeping this impression in mind compare it with the other surface. Fix your eyes on the other black square and once again you will get an impression of speed. Adjust the speed of the variable surface until the impressions are equal. Any questions?

The instructions for pursuit were as follows:

In equating the speeds, fix your eyes on a spot on the surface itself. Pick out one at the top of the window and follow it down until it disappears. Continue to do this until you get an impression of speed. Then do the same for the other surface, fixating and following down, and again you will get an impression of speed. Equate the two impressions. Do not count or use any similar procedure, but always use the impression obtained while following a spot. Any questions?

Before beginning the experiments, *O* was told to keep the back of his head against a headrest and was informed that he would be allowed 1 or 2 min. for the match. He then put on a pair of earphones transmitting a masking noise similar to that made by the motor of the two machines, thus making an auditory match impossible. *E* set the speed of the standard belt at 20 ft. per min. and that of the variable at some arbitrary value above or below this, alternating on successive trials. *O* then made his match. One practice trial was given prior to the first fixation judgment and another prior to the first pursuit judgment, the *O* being queried as to whether he could follow the instructions. The experiment was continued only when *E* was satisfied that *O* could use the specified mode of eye-behavior.

After each judgment the speed of the variable belt (also the standard) was measured by a tachometer. Twenty judgments were obtained from each *O*, 5 under each of the 4 conditions, or 10 with fixation and 10 with pursuit. Twenty-four college students were employed as *O*s.

RESULTS

The score for each *O* in Table I is the mean of all matches by *O* for each of the categories of the table. The difference between these means and the velocity of the standard is an error in feet per minute (ft./min.). This error score (without regard to sign) was used in the computation of the mean average errors and (with regard to sign) in the computation of the mean constant errors of Table I. The *N* of Table I is 24. The difference between 1.13 ft./min., the mean error for fixation and 1.25 ft./min.; the equivalent error for pursuit, was not statistically significant. There is

no evidence that one mode of observation is significantly more accurate than the other. Both yield judgments of considerable precision, the mean average errors being of the order of 5-6%. The constant errors for fixation and pursuit were not significantly different from zero by *t*-test.

Analyses of variance were run on the constant and average errors. Twenty error terms per *O* were used in each analysis. A space error (a

TABLE I
MEAN AND CONSTANT ERRORS IN FEET PER MINUTE FOR MATCHING
SPEED WITH TWO MODES OF OBSERVATION
(*N*=24)

	Mean of av. errors	SD	Mean of const. errors	SD
Fixation	1.13	0.46	+0.10	0.71
Pursuit	1.25	0.41	-0.12	0.73
Standard on Right	1.21	0.35	+0.25	0.82
Standard on Left	1.17	0.53	-0.22	0.93

tendency to see motion either more accurately or faster when looking in one direction than when looking in another) was significant at the 5% level.

DISCUSSION

A man who is trying to judge the speed of a moving automobile can do so in either of two ways: by fixating on the automobile itself (in which case there is a stationary retinal image of the automobile with a moving retinal image of the world) or by fixating on a telephone pole (in which case there is a moving retinal image of the automobile with a stationary retinal image of the world). It may be true that in the latter case the motions seem faster than the former, and that the two modes of observation are not equivalent. All *O*s in the present experiment agreed that the 'impressions of speed' under the two conditions were different. Nevertheless this experiment seems to show that one can *discriminate* speed just about as well in one way as the other. Untrained but carefully instructed *O*s can distinguish small differences in speed whether the eyes pursue or are fixated. In this important respect the two modes of observation are equivalent. The basis for this equivalence is an important problem for any theory of visual perception.⁴

SUMMARY

The accuracy with which the speed of two moving surfaces (when

⁴J. J. Gibson, The visual perception of objective motion and subjective movement, *Psychol. Rev.*, 61, 1954, 304-314.

seen behind identical rectangular windows) could be matched was determined for two different modes of observation: with fixation of the eyes and with pursuit by the eyes. Twenty-four *O*s made 20 judgments each, 5 per mode when the standard was on the right and 5 when on the left. The differences in errors between modes were not statistically significant. A space error was significant at the 5% level.

CONDITIONING OF FEAR: A FUNCTION OF THE DELAY OF REINFORCEMENT

By JOEL R. DAVITZ, DONALD J. MASON, and O. H. MOWRER, University of Illinois,
and PETER VIEK, Abbington Hospital, Philadelphia

The purpose of this study was to investigate the effect of delayed reinforcement on the acquisition of a conditioned emotional response. Specifically, the experiment was designed to determine the effect of paired presentation of a blinking light and an electric shock with various intervals of delay between the conditioned stimulus (CS) and the unconditioned stimulus (US).

The empirical fact that the strength of association is a function of the temporal proximity of the elements to be associated has long been recognized as a fundamental principle of learning. Critical reviews of the experimental evidence related to this phenomenon have been made by Spence and by Hilgard and Marquis.¹ In 1947, Spence summarized some of the existing research on the delay of reinforcement for instrumental learning. His review indicated that consistently shorter gradients of reinforcement had been obtained when the operation of secondary reinforcement had been minimized. He concluded as follows:

The fact that no learning occurs with delays as short as five seconds in the discrimination-situation suggests that the primary gradient of reinforcement, if there is such a thing at all, is much shorter than the 20 to 30 seconds inferred by Hull on the basis of the Perin studies. Indeed, it would not seem unreasonable to hypothesize that there is no primary gradient of reinforcement, but that all learning involving delay of the primary reward results from the action of *immediate* secondary reinforcement which develops in the situation. Factors affecting the development of this secondary reinforcement would then be considered as responsible for all gradients of learning found with delay of reward.²

A subsequent study by Grice has been interpreted as lending powerful support to Spence's contention that all learning under conditions of delay of reward is due to the effects of a step-by-step backward progression of secondary reinforcement.³ In Grice's study a very steep delay of reward gradient was obtained, with essentially no learning manifested for delays of reward exceeding 5 sec.

Many of the parameters influencing the gradient of reinforcement for visual discriminative learning and simple trial and error problem solving have been thoroughly investigated; however, relatively few experiments have been reported in which the extent of the gradient of reinforcement for avoidance- or fear-conditioning have been studied.⁴

* Received for publication December 14, 1955.

¹ K. W. Spence, The role of secondary reinforcement in delayed reward learning, *Psychol. Rev.*, 54, 1947, 1-8. E. R. Hilgard and D. G. Marquis, *Conditioning and Learning*, 1940, 158-175.

² Spence, *op. cit.*, 7.

³ G. R. Grice, The relation of secondary reinforcement to delayed reward in visual discrimination-learning, *J. exp. Psychol.*, 38, 1948, 1-17.

⁴ L. H. Warner, The association span of the white rat, *J. genet. Psychol.*, 41, 1932, 57-90; W. E. Kappauf and Harold Schlosberg, Conditioned responses in the white rat: III. Conditioning as a function of the length of delay, *J. genet. Psychol.*, 50, 1937, 27-45; L. J. Kamin, Traumatic avoidance learning: The effect of CS-US interval with a trace conditioning procedure, *J. comp. physiol. Psychol.*, 47, 1954, 65-72.

PRELIMINARY EXPERIMENT

A preliminary experiment by Mowrer and Vick was designed to measure the gradient of reinforcement for instrumental fear-conditioning. Twenty-eight Lashley-strain female rats, about 4 mo. of age, were reduced to and held at 85% of normal body weight (*i.e.* the weight of the animals when on full and unrestricted diet). These Ss were randomly assigned to seven groups of four animals each. The apparatus consisted of 7 box-like compartments, each approximately 12 in. square with a grid floor that could be electrified. A small electric lamp illuminated the interior of each compartment. Each S was placed in a compartment and, after 3 hr., a bit of wet mash (made of Purina Laboratory Chow) was offered to the animal on the end of a stick which was projected upward toward the animal from beneath the grid floor. One presentation of food and shock constituted a 'trial,' and only one trial was given per day for 15 days. Three hours after the daily trial S was returned to his home cage and fed sufficiently to hold his weight at the desired level.

On the first day, each of the 28 Ss immediately ate the food and the conveying stick was promptly withdrawn. They were then given 2 sec. of electric shock, 130-v. A.C., with a limiting resistance of 200,000 Ω , immediately or after 2, 5, 10, 30, 120, or 600 sec., depending upon the group to which the animal had been assigned in advance. Seven animals, one for each of the seven delay-intervals, were run concurrently, and the experiment as a whole consisted of four repetitions of this procedure. If S did not eat within 10 sec. after the food was presented, the food was withdrawn.

Presumably, as a result of the paired presentations of food and shock, presentation of the food would tend to elicit a conditioned emotional response. This emotional response was measured in terms of the number of times S entirely refrained from eating and in terms of the delay in eating shown by the Ss of each group. Fig. 1 represents graphically the results of the experiment. There was very considerable conditioning at 0, 2, 5, and 10 sec., and some conditioning at 30, and even 120 sec.

A reëxamination of the experimental procedure, however, revealed a possible artifact. To interpret the results in terms of a stimulus-trace (*i.e.* some persisting intraorganic after-effect produced by S having seen and eaten the food), one must make certain that there are not extraorganic stimuli which might serve to mediate the association between food and shock. In the procedure described above, there was one such possibility; namely, that particles of food might have remained in the animals' mouths, providing cues which would bridge the interval between food and shock. Therefore, the results were somewhat equivocal and were not independently published.

The present experiment was designed to eliminate the possibility of mediation between the CS and shock by extraorganic stimuli, thus permitting the results to be interpreted in terms of a stimulus-trace or the intraorganic after-effects of the presentation of the CS.

MAIN EXPERIMENT

Subjects. The Ss of this experiment were 56 male albino rats of the Sprague-Dawley strain. They were experimentally naïve and between 75 and 100 days old.

Apparatus. A multiple-unit apparatus, consisting of eight identical compartments, was used for training. Each compartment was 9 in. square with a grid floor that could be electrified. All walls were painted flat black and each compartment was

illuminated by a 7.5-w. light attached to the lid of the compartment. A mirror was placed under the floor of each compartment to permit observation of the rat in the compartment.

Training procedure. Training consisted of 20 trials in which the CS, a 2-sec. blinking light, was paired with a 3-sec. electric shock of 175 v. administered through 150,000 Ω resistance. Eight Ss, one at each CS-US interval, were trained concur-

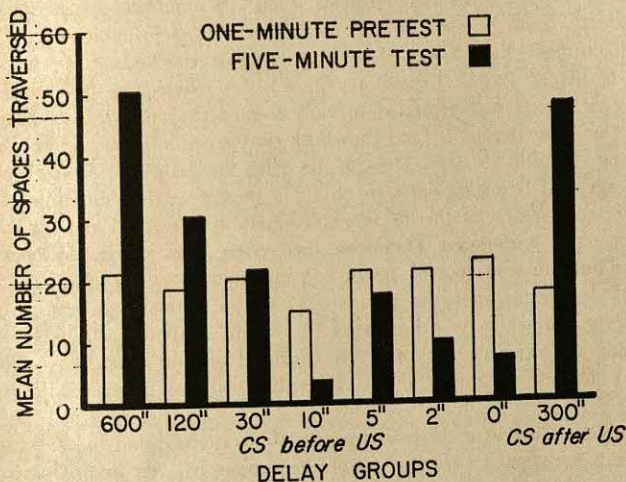


FIG. 1. TWO MEASURES OF INHIBITION OF EATING FOR SEVEN GROUPS OF ANIMALS WITH VARIOUS INTERVALS BETWEEN EATING AND SHOCK

rently. The CS, for different groups, was presented at intervals of 0, 2, 5, 10, 30, 120, and 600 sec. before the onset of the shock, and 300 sec. after the termination of the shock.

Each S received four training trials per day for five days. The first trial began 50 min. after S had been placed in the training apparatus. An interval of 50 min. was allowed between each of the 4 daily training trials; during this interval the animals remained in the apparatus. Fifty minutes after the last daily trial, S was returned to its home cage and fed. During training S was kept on approximately 20-hr. food deprivation to reduce defecation and to minimize the possibility of its shorting the circuit during the shock.

Testing procedure. The Ss were tested on the sixth day of the experiment. The test-procedure was designed to determine whether or not the CS at each of the eight CS-US intervals had acquired the capacity to evoke an emotional response. The behavioral index of the emotional response was based on the observation that rats tend to 'freeze' when they encounter a stimulus previously associated with punishment and from which they cannot escape. During the test-trial, the CS was presented and the amount of activity exhibited by the rat was measured. It was assumed that if the CS had acquired the capacity to evoke an emotional response, the rat would when the CS was presented, tend to 'freeze' and inhibit activity.

For the test-period, the partition between two separate compartments of the training apparatus was removed. Thus the test-compartment consisted of a box 18.75 in. long, 9 in. wide, and 9 in. high, with a grid floor and two 7.5 w. lights. The floor of the compartment was divided into 8 equal spaces, each 4.7 in. long and 4.5 in. wide. Activity was measured by counting the number of spaces traversed by the animals' right front paws during the test-period.

The procedure on the test-day consisted of a 1-min. pretest period of free exploration followed by a 5-min. test-period during which the CS was periodically administered. Each rat was placed in the test-compartment for 1 min. before the CS was initially presented. During the 5-min. test-trial, the blinking light was presented once every 30 sec. for the first minute and once every 60 sec. for the remaining 4 min.

Results. The 1-min. period of free exploration before the blinking light was administered, provides an estimate of the operant level, or initial activity level of

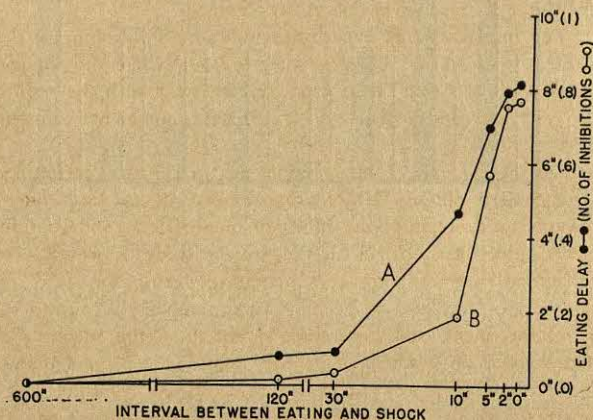


FIG. 2. AMOUNT OF ACTIVITY DISPLAYED BY THE EIGHT GROUPS OF ANIMALS DURING BOTH THE 1-MIN. PRETEST AND THE 5-MIN. TEST-PERIOD

each animal. The light bars in Fig. 2 present graphically the average number of spaces traversed by Ss in each of the eight delay groups during the 1-min. pretest. The variances of the eight groups were not homogeneous; therefore, nonparametric tests were used for the statistical analysis of the data. Testing the significance of the difference among the eight groups by an analysis of variance on ranks, a Chi-square for ranks of 9 with 7 degrees of freedom was obtained. The differences do not approach significance at the 10% level, hence the null hypothesis was accepted.

The dark bars in Fig. 2 represent the average number of spaces traversed by Ss in each group during the 5-min. test-period. It was assumed that there would be little or no conditioning when the CS was presented either 600 sec. before the shock or 300 sec. after the shock. This assumption was supported by a comparison of the activity exhibited by each of the groups at these CS-US intervals. A Mann-Whitney U of 22 ($p = 0.40$) was obtained, and the null hypothesis was accepted. In the statistical analysis of the data, each of the experimental groups was compared with

the 600-sec. control group. For the 0, 2, 5, 10, and 30-sec. groups the differences were significant at or beyond the 3.8% level of confidence. Although the difference between the 600-sec. and the 120-sec. groups was not statistically significant, the direction of the results suggests the possibility of some conditioning even at the 120-sec. interval. Four of the 7 animals in the 120-sec. group exhibited less activity than any of the 14 animals in the 600-sec. and 300-sec. control groups. It should be noted that only one of the control groups was utilized in the computation of these significance levels; therefore, these statistics may be regarded as conservative estimates of the differences between groups.

The test-trial activity of the 10-sec. group fell somewhat below that which might be expected on the basis of an interpolation of the gradient based on the other groups. As is indicated, however, in Fig. 2, the Ss in the 10-sec. group were also less active than the Ss in the other groups during the pretest-period. On the basis of this, it does not seem unreasonable to infer that these Ss may have been generally less active rats than those in the other groups. If this were the case, it would in part account for the slight discrepancy in the results obtained.

Discussion. The results of this experiment indicate conditioning at delays of reinforcement in excess of 30 sec. There is some indication that the obtained gradient may extend to beyond 120 sec. These data are consistent with the results of a preliminary study on trace conditioning which utilized a somewhat different response measure.

On the basis of the evidence from a number of experiments concerning the acquisition of instrumental responses, Spence proposed that all learning under conditions of delayed reward is a function of the effects of immediate secondary reinforcement.⁵ The experimental procedure utilized in the present study, however, precluded the possibility of secondary reinforcing cues operating during the delay period. The gradient obtained does not seem to be most parsimoniously accounted for in terms of secondary reinforcement, and casts some doubt upon the adequacy of Spence's hypothesis that there is no gradient of primary reinforcement, and that *all* learning under conditions of delay of reward is due to secondary reinforcement. Although there are superficial similarities among the various phenomena, and the various experimental procedures which have been utilized in studies of gradients of reinforcement,⁶ the present evidence does not indicate that all learning under conditions of delayed reinforcement can be conveniently attributed to the action of a single principle such as that of secondary reinforcement. Nor can one conclude, from the results of the present investigation, that the stimulus-trace is limited to a length of 30 sec. Even in the relatively simple conditioning situation utilized in the present study, it remains quite probable that the gradient could be further lengthened by properly manipulating the various experimental parameters such as the nature and intensity of the CS and US, the number of training trials, the sensitivity of the behavior measure, the degree of similarity between training and testing, and the species of animal utilized as an S.

SUMMARY

The purpose of the present study was that of investigating the acquisition of conditioned emotional responses under different conditions of delay of reinforcement.

⁵ Spence, *op. cit.*, 1-8.

⁶ Hilgard and Marquis, *op. cit.*, 158-175.

In the first study reported, the amount of inhibition of eating behavior was measured as a function of the degree of simultaneity between the presentation of food and electric shock. In the second study, animals were presented with a blinking light followed by an electric shock; different intervals between the CS and US were used with different groups of animals. The animals were then tested by observing the amount of 'freezing' behavior, or inhibition of activity which was produced by presentation of the blinking light alone. These procedures produced gradients far greater in length than have been previously found in visual discriminative studies. The theoretical implications of these results were discussed, with particular regard to the Spence hypothesis that all learning under conditions of delayed reward is due to immediate secondary reinforcement. It was concluded that this hypothesis lacks sufficient generality to encompass the present results.

TRANSFER AND RETENTION OF BIDIMENSIONAL COMPENSATORY TRACKING AFTER EXTENDED PRACTICE

By W. F. BATTIG, E. H. NAGEL, JAMES F. VOSS, and W. J. BROGDEN,
University of Wisconsin

Most laboratory experiments on motor learning have involved an amount of practice that is relatively small compared to the amount of practice apparently required for the acquisition of many skills of everyday life. It is possible there are variables that affect the acquisition of skill only at high levels of performance after extended practice. It is possible also that variables now known to be important in the early stages of acquisition may be relatively unimportant at the later stages. Furthermore, the nature of a skill may alter significantly as a function of the amount of practice over extensive temporal periods.

It was the purpose of the current research to examine some of these possibilities in the conduct of an experiment involving long-term practice of bidimensional compensatory tracking. The authors served as Ss in the experiment and practiced 10 trials per day for 100 days. Hypotheses about variables of importance in long-term practice were discussed at conferences. At the conclusion of the 100 days of practice, transfer-tests were made to determine the degree of independence of the skill thus acquired from the target-course and from the direction of control-movement involved during practice. Following additional practice to provide recovery of the skill after negative transfer, a lapse of approximately 8 mo. occurred before performance was measured again to test retention.

Procedure. The apparatus has been described in a previous study.¹ It consists of (a) a target-course generator, (b) a cathode-ray oscilloscope (CRO), (c) a stick-control, (d) two electronic integrators for scoring tracking-error separately for azimuth and elevation, (e) a clock for scoring time-on-target in azimuth and elevation, and (f) the necessary power-supplies and circuits for coupling of the components. S viewed the CRO-display through a rubber hood 9 in. long and manipulated the stick-control with his right hand. The target-scoring area was defined by a circle of 0.5 in.-diameter and the clock scored time-on-target when the $\frac{1}{8}$ -in. target-circle was within the 0.5-in. scoring-square. The maximal error-magnification of 0.48 in. of displacement of the target or error-marker on the CRO-face per degree of difference between the selsyns on the control-stick and course-generator was used to provide the maximal task-difficulty reported by Battig, Nagel, and

* Received for publication January 16, 1956. This research was supported in part by a grant from the Research Committee of the Graduate School from funds provided by the Wisconsin Alumni Research Foundation and in part by the United States Air Force under Contract No. AF33(038)-23294 monitored by the Armament Systems Personnel Research Laboratory, Air Force Personnel and Training Research Center, Lowry Air Force Base, Denver, Colorado.

¹W. F. Battig, E. H. Nagel, and W. J. Brogden. The effects of error-magnification and marker-size on bidimensional compensatory tracking, this JOURNAL, 68, 1955, 585-594.

Brogden.² There was direct stimulus-response compatibility throughout the experiment except for the transfer-trials of practice periods 107-110 when there was reversal. The standard condition associated right-left and up-down movement of the error-marker with right-left and backward-forward movement of the stick.

Each of the 4 Ss received 10 trials of the 1-min. target-course during each of the sessions. The inter-trial interval was 20 sec. except for a rest of 1 min. between the fifth and sixth trials. Since the Ss were familiar with the apparatus, knew the plan of the experiment, and had previous practice in tracking, no special instructions were given. Each S saw his scores daily and at least once a week, observed the plot of his learning curves for azimuth-integrator score, elevation-integrator score, and time-on-target in azimuth and elevation. The procedure and reference-measures for converting integrator-readings in volts to error of stick-position in degrees are the same as those used in the earlier study.³

All Ss completed 100 standard practice-sessions. For Sessions 101 and 102, the course-cam was operated in the reverse direction on each trial. Standard practice sessions were given on Sessions 103-106, but on Sessions 107-110 the direction of control-movement relative to target-movement was reversed for both azimuth and elevation. Sessions 111-114 provided practice under the standard conditions. Approximately 8 mo. later (223 days), 4 standard practice sessions (338-341 days after starting the study) were completed on 3 of the original 4 Ss to provide measures of retention.

Results. The results for the several conditions of the experiment are presented for time-on-target, integrated azimuth-error, and integrated elevation-error in Fig. 1. Performance in the first 100 practice sessions represents the learning curves. All measures of performance give a similar picture of acquisition with the time-on-target curve being the inverse of the two error-curves. The slowness of acquisition is evidence of the relatively great difficulty of the tracking task. An asymptotic level of performance was reached by approximately the eightieth session.

The analysis-of-variance technique described by Edwards for repeated measures on the same Ss was used for testing the significance of the results on acquisition, and for all subsequent tests of transfer and retention.⁴ In all cases, the 5% level of confidence was the criterion for statistical significance. Statistically significant acquisition was obtained for all measures of tracking between the first 10 and last 10 practice sessions. No statistically significant improvement in performance occurred after the eightieth session.

The weekly conferences produced a number of hypotheses. Many did not stand up under observations made during later stages of practice and many were of such character that they could not be investigated experimentally under the conditions of the present study. It was a common observation of all Ss, however, that as acquisition proceeded, there appeared to be less and less information available about the target-course. None of the Ss could verbalize the target-course in any way nor believed that he had learned the course. Thus, the hypothesis was formulated that with bidimensional compensatory tracking there is little learning of the course as such. The learning consists primarily in the acquisition of skill in making adjustive

² *Ibid.*, 585-594.

³ *Ibid.*, 585-594.

⁴ A. L. Edwards, *Experimental Design in Psychological Research*, 1950, 1-432.

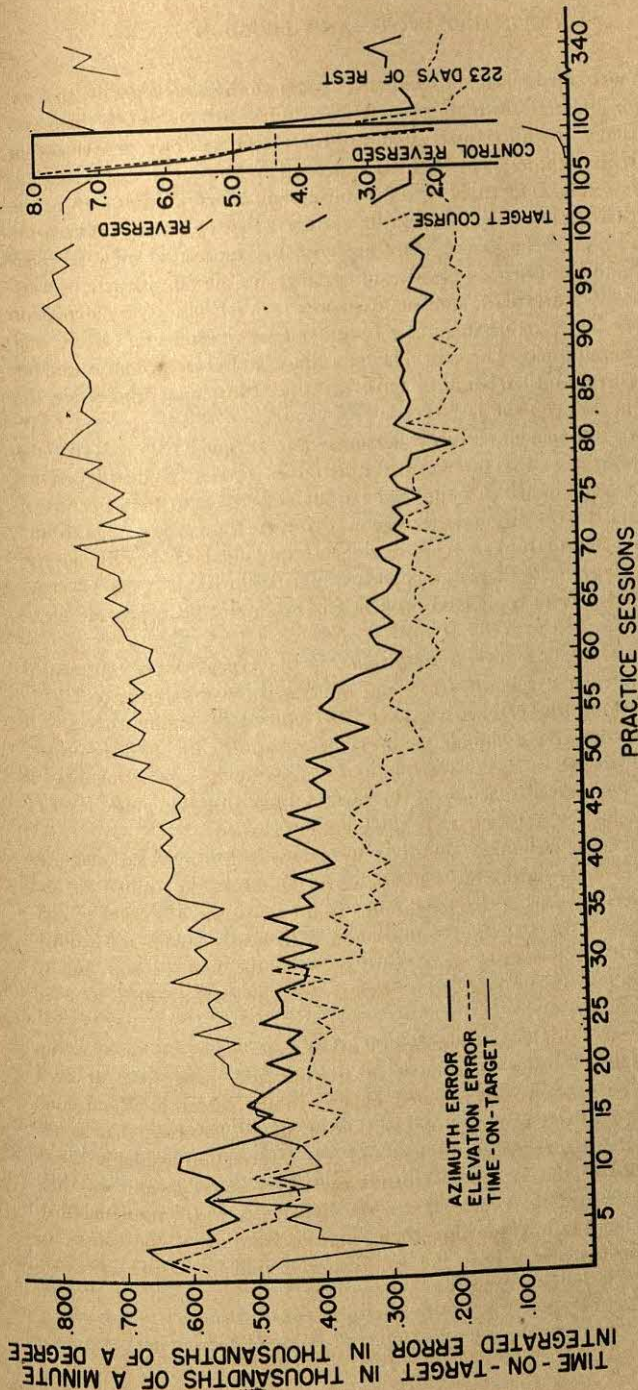


FIG. 1. ACQUISITION, NEGATIVE TRANSFER, AND RETENTION IN BIDIMENSIONAL COMPENSATORY TRACKING

The curves for azimuth-error, elevation-error, and time-on-target are plots of the group-mean for the 10 trials of each practice session, except that the curves for retention are based on only three Ss. Reversal in the direction of control-movement produced so much negative transfer that the ordinate-scale for integrated error had to be reduced by a factor of 10 in order to plot scores for azimuth- and elevation-error. The solid and broken lines parallel to the abscissa for the section of the graph giving the results for reversal of direction of control-movement represent the integrated error-scores for azimuth and elevation respectively produced by the course-cam with no operator present.

movements of the control so quickly and precisely that minimal error is displayed. If this hypothesis holds, transfer to other target-courses will be positive and high. It also follows that, if the directions of control-movement are reversed, there will be negative transfer.

The adequacy of experimental design for measuring transfer in practice sessions beyond the 100th is seriously limited by the small N , the long interval between initial measures of performance and the transfer-tests, and the absence of control groups. Even so, the results do provide support for the proposed hypotheses. For Sessions 101 and 102, when the course-cam was operated in reverse, all measures show performance inferior to that during terminal acquisition, but superior to performance early in practice. There are statistically significant differences between performance on the last 10 trials of acquisition and the first 10 trials of reversed course for all measures, on the last 10 trials of acquisition and the second 10 trials of reversed course for azimuth-error and time-on-target, and between performance on the first 10 practice sessions and the first 10 trials of reversed course for azimuth- and elevation-error. Although a plot of individual trials shows some evidence of learning under the reversed-course condition, there are no statistically significant effects of practice on performance. The plot of individual trials also shows an apparent initial transfer-effect for the error-measures from the reversed-course condition to the standard condition of tracking. There are, however, no statistically significant differences in performance from that on the last 10 trials of acquisition.

Reversal in direction of control-movement resulted in serious disruption of performance and high negative transfer. The scale for integrated error of Fig. 1 must be reduced by a factor of 10 in order to plot performance. The parallel lines give the integrated error-scores produced by the course-cam with the stick fixed in the vertical plane. The time-on-target score for the same condition is 0.001 min. During the initial trials, the target was off the scope most of the time and the S s had great difficulty in maintaining any control. When the target did appear on the scope, responses nearly always put it off again. Performance was poorer than in the no-operator condition. After 20 trials of practice, performance improved to approximately the no-operator level. The target was on the scope much of the time during each trial, but the error was large. Even after 40 trials of practice, performance was considerably poorer than the initial level of performance under standard conditions. The results of F -tests support all of the above conclusions, except that the differences in performance on azimuth and elevation for the last 10 trials of reversed control-direction and the first 10 sessions of acquisition are not significant. Statistically significant improvement in performance occurred over the 40 trials of the condition of reversed control-direction.

There was significant negative transfer from the condition of reversed controls to the first few trials of performance on the standard condition. Recovery was complete by the second practice session following the condition of reversed controls and performance thereafter on all measures was not significantly different from that on the last 10 trials of acquisition.

Performance after 223 days of rest shows a high degree of retention for the three S s who were available for these measures. The initial trials of retention give measures of performance that are numerically poorer than performance late in acquisition and just prior to the retention-interval. None of the session-means is sig-

nificantly different from the appropriate means for sessions late in acquisition nor for sessions just prior to the retention-interval. Other analyses of variance of performance as a function of all trials during the first two retention-sessions (Days 338 and 339) show no statistically significant practice-effect.

Discussion. The relatively great difficulty of the tracking task is probably of major importance in determining the results obtained on acquisition and transfer. Difficulty, in turn, is primarily a function of the large magnification of error, 0.48-in. displacement of the error-marker on the face of the scope per degree of difference between the selsyns on the control-stick and the course-generator. When the level of difficulty is expressed in terms of amplitude of hand-movement, an error of greater than 0.0613 in. (0.517°) places the error-marker outside the time-on-target scoring zone, and an error of 0.633 in. (5.169°) puts the error-marker off the face of the scope. Thus, the level of difficulty is near the maximum at which tracking is feasible. The results on acquisition of the present study taken with those obtained earlier suggest that an asymptotic level in performance of bidimensional tracking would be obtained with progressively smaller amounts of practice down to a magnification of error of 0.24 in. per degree at which speed of acquisition would be maximal.⁵

The results of the transfer-tests support the hypothesis that what is learned primarily in the tracking task is precision in making adjustive movements of the control. Within the limits of the experimental design, there is evidence for positive transfer to a different target-course and for high negative transfer when the direction of control-movement is reversed. The latter result conforms to the principle proposed by Gagné, Baker, and Foster that complete or partial reversal is the only type of task-alteration that leads to negative transfer.⁶ It is also in agreement with the negative transfer obtained by Lewis, McAllister, and Adams with reversal on a motor task.⁷ Reversal of control-direction in a motor task also involves alteration of S-R compatibility, which, as Fitts and Seeger point out, requires the use of additional control procedures in transfer-studies.⁸ Since the directions of control movement used on test-periods 109-112 probably represent the least degree of S-R compatibility and the directions of control movement used during acquisition represent a very high degree of S-R compatibility, the difference in performance is a function both of reversal and of change in S-R compatibility. With adequate control of compatibility, the amount of negative transfer would probably be less than that measured from the reference of initial performance in the present study. The large negative transfer thus obtained is a function of difference in S-R compatibility, reversal in direction, and amount of practice in original learning. Amount of negative

⁵ Battig, Nagel, and Brogden, *op. cit.*, 585-594.

⁶ R. M. Gagné, K. E. Baker, and Harriett Foster, On the relation between similarity and transfer of training in the learning of discriminative motor tasks, *Psychol. Rev.*, 57, 1950, 67-79.

⁷ Don Lewis, D. E. McAllister, and J. A. Adams, Facilitation and interference in performance on the Modified Mashburn apparatus: I. The effects of varying the amount of original learning, *J. exp. Psychol.*, 41, 1951, 247-260; McAllister and Lewis, II. The effects of varying amount of interpolated learning, *ibid.*, 41, 1951, 356-363.

⁸ P. M. Fitts and C. M. Seeger, S-R compatibility: Spatial characteristics of stimulus and response codes, *J. exp. Psychol.*, 46, 1953, 199-210.

transfer under reversal was found to be directly related to amount of original learning by Lewis, McAllister, and Adams.^o

Summary. Acquisition of bidimensional tracking with high error-magnification was measured by integrated error for azimuth and elevation and by time-on-target with 4 Ss for 100 practice sessions. Asymptotic performance for all measures was obtained by approximately the eightieth sessions. The error-curves show the same form and the time-on-target curve is the inverse of the error-curves. Positive transfer was found for the new course provided by reversing the course-cam and high negative transfer was found for reversal of control-direction. Retention of the skill by 3 of the Ss after 227 days of no practice was found to be very high.

^o Lewis, McAllister, and Adams, *op. cit.*, 247-260.

SCALING LINE-LENGTHS WITH A MODIFICATION OF THE CONSTANT-SUM METHOD

By KATHERINE E. BAKER and FRANK J. DUDEK, University of Nebraska

The procedure generally used in the constant-sum method is patterned after the psychophysical method of paired comparisons.¹ Instead of two categories of response for all possible pairings of a set of stimuli, a constant sum of points, usually 100, is divided between the members of each pair to express relative magnitude. Such a procedure provides both direct and indirect estimates of how many times greater any stimulus is judged to be in comparison with others. In determining scale-values for the stimuli, the full array of judgments is utilized, and combining both types of estimate presumably gives more reliable and valid ratios upon which to base scale-values.

Recent studies in our laboratory point to the likelihood that different judgmental biases operate when the pairs judged are similar in magnitude than when they are quite dissimilar.² If the biases noted are generally characteristic of the types of judgement made in the constant-sum method, the operational meanings of resultant scale-values are not so clear-cut as they at first might appear. It would be important to recognize that a direct estimate involving dissimilar stimuli might reflect certain biasing factors, while an indirect estimate of the same ratio from other pairs with less dissimilarity in stimuli may be influenced by factors other than those operative in the direct judgments. Thus, the average of both direct and indirect estimates may be quite different from the mean of a comparable number of direct estimates only. Certain implications of scale-values obtained by the constant-sum method may not hold when direct and indirect estimates do not come from a common population of estimates.

Ratio-determinations with clearer operational meanings might be obtained through repeated measurements of the direct estimates alone as could be accomplished by using the method of constant stimulus as the model instead of the method of paired-comparisons. With the chosen 'unit'-stimulus as the standard, point-divisions could be made for each possible pairing of the standard with the other comparison-stimuli of the set. Experimental time would be saved by restricting judgments to those particular stimulus-pairs for which information is desired, and computation of scale-values from judgmental data would be considerably simplified over what is required with the design of the method of paired-comparison.

The present experiment was devised to investigate the scaling of line-lengths

* Received for publication February 15, 1956. We are indebted to the University Research Council, University of Nebraska, for financial assistance in this research.

¹ Milton Metfessel, A proposal for quantitative reporting of comparative judgments, *J. Psychol.*, 24, 1947, 229-235; A. L. Comrey, A proposed method for absolute ratio scaling, *Psychometrika*, 15, 1950, 317-325; J. P. Guilford, *Psychometric Methods* (2nd ed.), 1954, 208-220.

² Trygg Engen, An evaluation of a method for developing ratio scales, this JOURNAL, 69, 1956, 92-95; K. E. Baker and F. J. Dudek, Weight scales from ratio judgments and comparisons of existent weight-scales, *J. exp. Psychol.*, 50, 1955, 293-308.

using such a modification of the constant-sum method. The particular dimension was chosen to permit comparison of the scale-values obtained with those found by Engen for the same stimulus-materials but with the design of the method of paired-comparison.³ Special attention was given to the consistency of scale-values developed with the modified technique.

METHOD

Subjects. Two groups of volunteers from undergraduate psychology courses were employed as Ss. Group A consisted of 49 Ss, 29 men and 20 women, who made a single series of judgments in one experimental session. Group B was made up of 26 different Ss, 19 men and 7 women, who repeated their judgments one month after an initial experimental session. This group will be referred to as Group B₁ for the first session and Group B₂ for the second.

Stimulus-materials. This experiment employed ten 2 × 2-in. photographic slides containing white lines, one above the other, on a black ground. The slides paired

TABLE I

SCALE-VALUES FOR STIMULUS-LINES AS DETERMINED BY DIFFERENT
EXPERIMENTAL GROUPS AND BY PHYSICAL MEASUREMENT

Groups	No. Ss	Stimulus-lines										
		A	B	C	D	E	F	G	H	I	J	K
A	49	1.00	1.35	1.74	2.31	3.22	4.19	6.04	7.59	10.86	17.39	24.99
B ₁	26	1.00	1.35	1.77	2.41	3.13	4.50	6.46	7.74	10.69	13.85	19.64
B ₂	26	1.00	1.31	1.77	2.41	3.31	4.14	6.06	7.23	10.56	16.05	25.40
Physical measure		1.00	1.33	2.00	3.00	4.00	6.00	9.00	12.00	18.00	27.00	36.00

the shortest line, *A*, with each of the 10 longer lines, *B* to *K*. Physical measurement of the 11 lines gave the following ratios: $A/B = 1.3$, $A/C = 2.0$, $A/D = 3.0$, $A/E = 4.0$, $A/F = 6.0$, $A/G = 9.0$, $A/H = 12.0$, $A/I = 18.0$, $A/J = 27.0$, $A/K = 36.0$.

Procedure. The Ss were tested in small groups with the slides projected onto a large screen in a 25-ft.-long, darkened room. The size of the projected lines and seating positions were so arranged that the stimulus-lines subtended very nearly the same visual angles as in Engen's experiment. The Ss were seated within a few feet of the projector to keep the line of regard nearly perpendicular to the screen.

In the first experimental session, the task of dividing 100 points between each pair of stimuli was explained to the Ss and the implications of several hypothetical point-assignments were indicated. The use of the data-sheets provided was explained, and the Ss were cautioned against influencing each others' judgments by talking during the session. Each of the 10 slides was shown once with line-length *A* as the top and once as the bottom member of the pair, making a total of 20 judgments. The 20 pairs were presented in a predetermined random order and each was exposed for about 30 sec., with a 10-sec. warning signal before the exposed pair was removed from the projector.

RESULTS

Points assigned to line *A* in comparison with each other line were totaled and the ratio of the sum of all possible points to the total for *A* was calculated for each comparison-line. Table I presents these ratios which are the psychological scale-values with line-length *A* as the unit.

³ Engen, *op. cit.*, 92-95.

Fig. 1 shows a plot of psychological scale-values as a function of physical values. The diagonal straight line is the function expected if there were a one-to-one correspondence between the two scales. For comparison purposes, Engen's results for 25 Ss are presented in two ways: the scale-values obtained using the full array of judgments of all possible pairings of the stimuli as in the paired-comparison

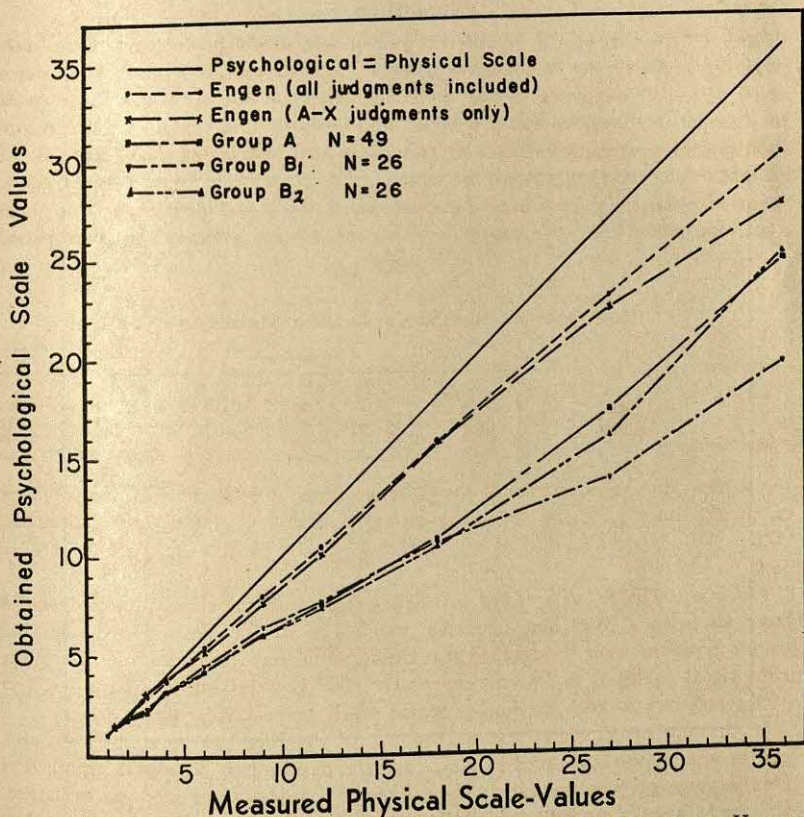


FIG. 1. THE RELATION BETWEEN PSYCHOLOGICAL AND PHYSICAL SCALE-VALUES

design, and the scale-values from just the *A/X*, or direct estimates, abstracted from these data.

The results show that for all three sets of our data the psychological scale-values are underestimates of the physical values, with the exception of very slight overestimations for line *B* for Groups *A* and *B*₁. The extent of underestimation increases with the magnitude of stimulus-ratios, *i.e.* the dissimilarity of stimuli judged. The findings are highly reliable. Groups *B*₁ and *B*₂ show that retest scale-values agree very closely with those of the first session. The greatest discrepancies appear with the two longest line-lengths, but these changes represent decreases of only about two points in the mean number of points assigned line *A* in the two ses-

sions. Further, scale values from independent Groups A and B₁ and B₂ also are consistent.

The present results give consistently lower scale-values, with the exception of line B, than Engen obtained using the paired-comparison model combining direct and indirect estimates of ratios. The direct (A/X) estimates abstracted from his data are generally lower than the values he reported for the averaged estimates, but they are not as low as the scale-values obtained with the modified technique employed in this experiment.

Of interest, too, is the variability of point-assignments for the various line-pairs. Fig. 2 shows two different ways of looking at variability (obtained for Group A) in

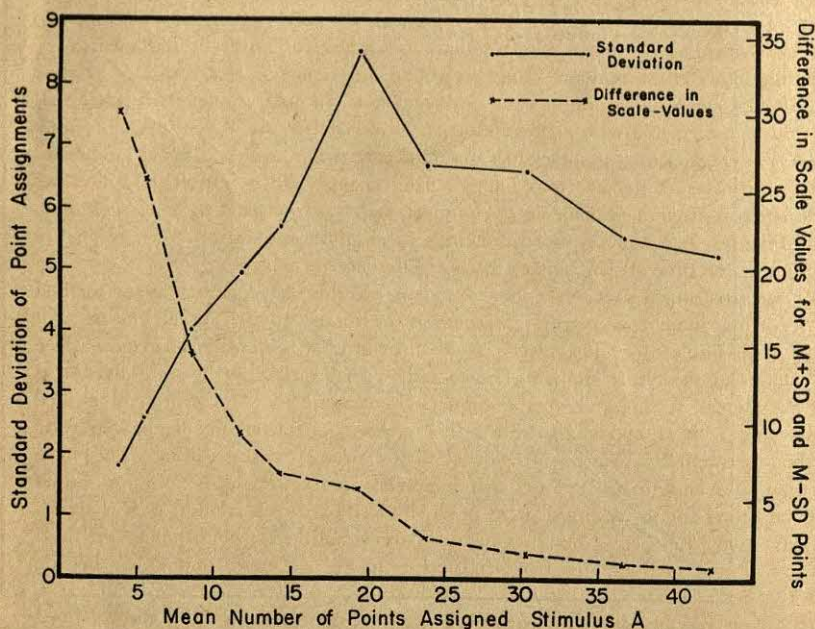


FIG. 2. VARIABILITY OF POINT-ASSIGNMENTS AS A FUNCTION OF THE NUMBER OF POINTS ASSIGNED TO STIMULUS A (GROUP A)

the present experimental situation; in terms of *SDs* for point-assignments and in terms of the variability in scale-values produced by the variabilities in points. The solid-line, to be referred to the left-hand ordinate, shows that the *SDs* of points are largest for medium degrees of similarity (line A given 20-30 points) and smallest for dissimilar pairs (A given few points). To show the extent of variability in scale-values, computations were made of the scale-value which would have resulted had the point-assignment been one *SD* higher than the *M*, and the value had the point-assignment been one *SD* less than the *M*, and then the difference between these values was taken as the measure of scale-value variability. The dashed line in Fig. 2, to be referred to the right hand ordinate, shows that variability in scale-values is

high for dissimilar stimuli, where *SDs* are small and rapidly decreases with increasing similarity of stimuli in spite of the larger *SDs*.

DISCUSSION

At least for judgments of line-length, a model after the method of constant stimulus appears to result in lower scale-values than does the previously used technique employing a paired-comparison model. Such scale-values may be reliably determined with repeated measurements of a few stimulus-pairs and computational procedures are easily carried out. Comparable data abstracted from Engen's previous works also give lower scale-values than when the method of paired-comparison was used. It is reasonable to believe that the exclusion of indirect estimates of ratios as determinants of scale-values is to some degree responsible for the lowering of scale-values in our experiment compared to Engen's findings for the full array of judgments.

Perhaps this effect is due to some extent to differences in perception of relative magnitudes when one stimulus-value is common to all pairs rather than having all stimulus-values paired with all others. Michels and Helson show the effects of variations in aggregates of stimulus-values upon the adaptation-level (*AL*) and judgment.⁴ Direct estimates abstracted from Engen's data do not yield as different scale-values as when a restriction is made in the number and types of pairs of lines presented. If differences in *S*-samples and differences in absolute size of stimulus-objects are negligible, the present data suggest such a difference in perception.

As we have noted previously, however, it is possible that artifacts of the method of reporting judgments may be erroneously attributed to perceptual effects.⁵ For pairs of dissimilar stimulus-objects, small differences in point-divisions mean large differences in the ratios implied. The data show that variability in point-division is not very great at this end of the similarity-continuum, but variability in implied scale-value is at a maximum. When this inherent characteristic of the reporting method is coupled with a likely tendency to shy away from the 'ceiling' 1.99 point-division, it seems to be expected that increasing underestimates of physical values will be found for increasing degrees of dissimilarity between stimuli, as was shown in this study. Restricting *S*'s judgments to direct estimates would show up these effects without confounding from indirect estimates which include other biasing factors.

Although the operational meaning of scale-values determined by the modified technique is clearer than when the paired-comparisons model is used, it does not follow that such scale-values are any more valid as predictors. While *A/B*, *A/C*, *A/D*, and so on, are given directly by the new technique, *B/C*, *B/D*, *C/D*, and so on, are only implied and may not be the same as would be obtained by direct estimates. Thus, the modified technique, though convenient, may have its own problematic aspects and should not be considered as the ultimate in refinement of the constant-sum method. Variations in specific techniques in the constant-sum method produce variations in obtained scale-values, and the question of choice among techniques must await experimental demonstrations of predictive power and internal consistency of the scales developed.

⁴ W. C. Michels and Harry Helson, A reconciliation of the *VEG* scale with Fechner's law, this *JOURNAL*, 67, 1954, 677-683.

⁵ Baker and Dudek, *op. cit.*, 293-308.

SUMMARY

Since the constant-sum method with the paired-comparison model does not provide scales with clear operational meaning, line lengths were scaled using a modified technique with the method of constant stimulus as the model. One hundred points were divided between members of each pairing of a short 'standard' line with each of 10 longer lines. The modified technique yielded highly reliable but different scale-values from those previously obtained with the constant-sum method. Increasingly large underestimates of physically measured values appeared with increasing dissimilarity between stimuli judged. Variability of scale-values was shown to be a complex function of variability of point assignments.

PREFERENCES OF MONKEYS FOR OBJECTS OTHER THAN FOOD

By JAMES B. CHO and ROGER T. DAVIS, University of South Dakota

The present investigation was designed to study the preferences of monkeys for selected lures that were not food. The possibility of such an investigation was suggested by the finding of McDowell that there was preferential manipulation of parts of the living cages.¹ McDowell, however, was not able to ascertain what aspect of the particular objects in and about the cage may have contributed to the monkey's preferences. In the present study, objects were compared which varied in shape and in the nature of the material from which they were formed.

Relative preferences for different foods have long been known. Tinklepaugh demonstrated such preferences by showing that the substitution of a less-preferred food in a delayed reaction-experiment disrupted the animal's behavior.² Food preferences of monkeys were found, in later experiments by Harlow and his coworkers, to be related systematically both to the quantity and the quality of the food-reward.³

Harlow has also reported experimental evidence to show that monkeys persistently manipulate mechanical puzzles, open doors, and learn to make discriminations without the usual reinforcement with food or water.⁴ He has labeled this behavior evidence for a *manipulation drive*, and has pointed out many differences between homeostatic and externally aroused drives.⁵

Method. The experimenters employed special procedures that had been shown to be appropriate for studying the responses of monkeys to lures that were not food. Trials were spaced 5 hr. apart each day because Davis and McDowell had noted that puzzle manipulation not rewarded with food was much more vigorous with 24 hr. between trials than in earlier studies with massed trial presentation.⁶ A reasonable size for the objects was indicated by the persistent manipulation by monkeys of wood fragments, hog rings and small loose pieces of wire noted in McDowell's extensive naturalistic observations.⁷ Selection of the materials used in the construction of the objects was determined by need to make objects of different material as alike in shape as possible.

* Received for publication July 27, 1955. From the Radiobiological Laboratory, University of Texas. This investigation was supported in part by Grant M-530 from the National Institutes of Health.

¹ A. A. McDowell, A normative study of the behavior of the caged macaque, *Amer. Psychologist*, 8, 1953, 395-396.

² O. L. Tinklepaugh, An experimental study of representative factors in monkeys, *J. comp. Psychol.*, 8, 1928, 197-236.

³ J. C. Fay, J. D. Miller, and H. F. Harlow, Incentive size, food deprivation, and food preferences, *J. comp. physiol. Psychol.*, 46, 1953, 13-15.

⁴ Harlow, Motivating factors underlying learning, in *Learning Theory, Personality Theory, and Clinical Research, The Kentucky Symposium*, 1954, 36-53.

⁵ Harlow, Motivation as a factor in the acquisition of new responses, in M. R. Jones (ed.), *Current Theory and Research in Motivation*, 1953, 24-49.

⁶ R. T. Davis and A. A. McDowell, Manipulation of a 6-unit mechanical puzzle by monkeys under spaced practice, *Proc. S. Dak. Acad. Sci.*, 32, 1953, 143-146.

⁷ R. T. Davis, A. A. McDowell, and J. P. Steele, Performance of rhesus monkeys on related laboratory tasks presented before and after a large single dose of whole body x-radiation, *J. comp. physiol. Psychol.*, 49, 1956, 20-26.

Monkeys usually simultaneously sight and reach for an object, a tendency that made it possible for the experimenters to insure sequential choices of the objects by designing a holder for the lures consisting of long parallel tubes.

Subjects. The Ss were 15 laboratory-adapted rhesus monkeys. They were housed in pairs in 3 ft. $\times$ 3 ft. $\times$ 3 ft. cages constructed of steel bars. Four months before

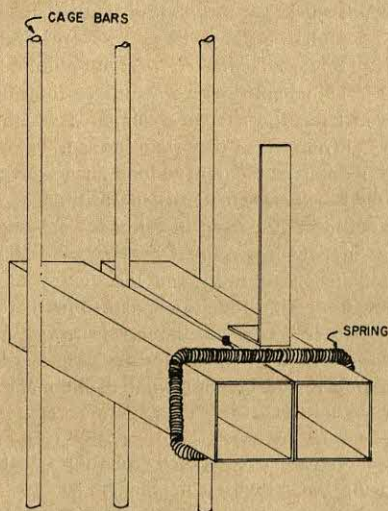


FIG. 1. STIMULUS HOLDING DEVICE

The paired hinged tubes are shown on either side of a cage-bar. The apparatus was pushed into the cage until the L-shaped strap was in the plane of the bars.

the present experiment began, nine of the Ss had been irradiated with X-rays (400 r whole-body irradiation). The six Ss, not irradiated, comprised a control group. By the beginning of the present experiment, all measured differences between the puzzle or problem-solving behavior of the animals in the irradiated and control groups had disappeared.

Apparatus. The apparatus used in this experiment is shown in Fig. 1. It consisted of a stimulus-holding device and 12 stimulus-materials. The holder for the lures consisted of two rectangular sheet steel tubes. Each tube measured 2 in. $\times$ 2 in. $\times$ 6.75 in. The two tubes were attached to each other at one end by means of a flat metal band and were held tightly one against the other by means of a screen-door spring that encircled both tubes 1.5 in. from the attached ends. An L-shaped steel strap was welded on to the top of one of the tubes.

The holder for the lures was mounted on the cage door by forcing the tubes apart against the action of the spring and by placing one tube on either side of the center vertical bar of the cage. The device was steadied by inserting the L-shaped strap under a piece of wire mesh which spanned the bars.

Four types of material—wood, paper, cloth, and metal—were fashioned into three shapes: elongate, flat, and solid. Swab sticks, tightly twisted tissue paper,

common cotton twine, and 14-gauge wire were used to make elongate wooden, paper, cloth, and metal lures, respectively. Bleached muslin, 16-lb. typewriter paper, galvanized tin, and cottonwood veneer were used to make flat cloth, paper, metal, and wooden lures, respectively. The solid cloth and paper lures were constructed by wetting, rolling, and drying bleached muslin and tissue paper. The resulting hard cloth and paper dowels were then cut into proper lengths. Steel and wooden dowelings were cut to make appropriate solid metal and solid wooden lures. The elongate lures measured 1 in. long and approximately 0.03 in. to 0.06 in. in diameter. The solid materials measured 0.5 in. long, and 0.5 in. in diameter.

Procedure. When each of the 12 different kinds of lures are paired together, two at a time, 66 possible pairings of stimulus-objects can be made, and since each material was also paired with itself, 12 additional, or 78 total, pairings were possible. By reversing the lateral order of each pairing, 156 pairings are produced. Each *S* was presented with three stimulus-pairings a day (7 A.M., 12 N., and 5 P.M.) for 52 days. The order of presentation of all 156 pairings of stimuli was individually randomized for each *S*.

A single trial was conducted in the following manner. The cage-mate of the animal to be studied was removed from the living cage and stored in a carrying cage. The choice-apparatus was inserted between the cage bars. One stimulus-object was placed at the end of each tube and two stop-watches were started. When the monkey removed one of the lures from a tube the experimenter stopped one of the stop watches, and when the animal removed the other object from the other tube the experimenter stopped the second stop-watch. The time taken by each animal to remove each object was recorded. When an animal did not remove either object within 60 sec., the trial was terminated and a failure was indicated on the data sheet.

Results. Throughout training the animals maintained a high level of responding. Stimulus-selections were typically successive, and both of the offered lures were usually selected within the 60-sec. period of time allotted per trial. In the first session of training every animal responded by removing one object of an object-pair within the time allotted, and throughout the remaining 155 experimental sessions the *Ss* selected at least one stimulus-object out of every pair on 99.7% of the trials. Furthermore, the *Ss* removed both stimulus-objects in a pair in 97% of the trials. Simultaneous choice was infrequent, and occurred on only seven of the 2350 trials given.

The average time taken by an *S* to remove each type of lure is shown in Fig. 2. The *Ss* responded with the longest latencies to elongate metal (7), elongate cloth (1), and flat metal materials (8); with the shortest latencies to solid cloth (2), solid wood (12), and solid paper (6). The three last named stimulus-materials were each chosen first significantly more times than any of the other materials. Table I summarizes the number of instances in which *Ss*' latencies of responding to one stimulus-object were longer or shorter than the latencies of his responding to other stimulus-materials. These results show that generally more animals have shorter latency of response to elongate materials (1, 4, 7, 10) than to solid materials (3, 6, 9, 12). Co-varying with this tendency, however, is the low preference (long latency of response) for metal materials (7, 8, 9). No differences were found between the mean latencies nor the rank order of the latencies of the re-

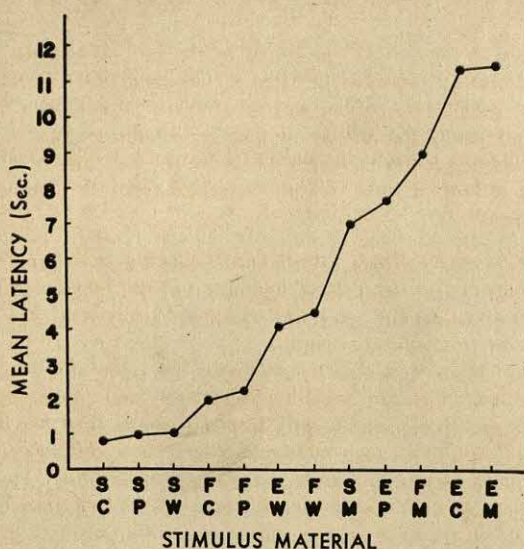


FIG. 2. PREFERENCES OF MONKEYS FOR VARIOUS, INEDIBLE STIMULUS-OBJECTS

S = solid cylinders

F = flat objects

E = elongate objects

C = cloth objects

P = paper objects

W = wooden objects

M = metallic objects

TABLE I

NUMBER OF ANIMALS RESPONDING TO THE STIMULUS-OBJECTS IN THE SIDE STUBS WITH A LONGER LATENCY THAN TO THE STIMULUS-OBJECTS IN THE CAPTIONS

(N=15)

Stimulus-objects		Cloth			Paper			Metal			Wood			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	
Cloth	Elongate	(1)	—	15	15	14	15	15	9	7	12	15	13	14
	Flat	(2)	(0)	—	13	(1)	8	13	(1)	(1)	5	4	3	12
	Solid	(3)	(0)	2	—	(0)	(0)	7	(0)	(0)	(1)	(0)	(0)	6
Paper	Elongate	(4)	(1)	14	15	—	14	15	3	7	10	8	10	13
	Flat	(5)	(0)	7	15	(1)	—	12	(0)	(0)	6	3	3	11
	Solid	(6)	(0)	2	8	(0)	3	—	(0)	(0)	2	(0)	(0)	8
Metal	Elongate	(7)	6	14	15	12	15	15	—	9	12	14	14	15
	Flat	(8)	8	14	15	8	15	15	6	—	13	11	13	15
	Solid	(9)	3	10	14	5	9	13	3	2	—	7	6	12
Wood	Elongate	(10)	(0)	11	15	7	12	15	(1)	4	8	—	6	13
	Flat	(11)	2	12	15	5	12	15	(1)	2	9	9	—	14
	Solid	(12)	(1)	3	9	2	4	7	(0)	(0)	3	2	(1)	—

Boldface figures indicate a significantly greater number of Ss responding with a longer latency (using sign test).
 Parenthetical figures indicate significantly fewer Ss responding with a shorter latency (using sign test).

sponses of animals previously irradiated with X-rays and the animals in the non-irradiated control group.

The performance of the animals, measured by the lure chosen first in a pairing, parallels the performance measured by latency. The rank order correlation is 0.96 between the ranks computed according to total response latency for a given material and the ranks representing the number of times a stimulus-material is selected first in a pairing. In addition to the consistency between the measures of first choice and latency, the ranks of both of these measures remained relatively constant throughout the entire experiment.

Discussion and summary. Three salient factors relating to manipulative drive of monkeys were contained in the present findings: (1) the particular lures chosen; (2) the persistence and rapidity of responses made by monkeys; and (3) the uniformity of preference from animal to animal.

Generally, metal objects were chosen significantly less frequently than paper, wood, or cloth, but the latter three do not differ consistently and significantly from each other. Solid lures are chosen significantly more frequently than flat lures, and flat lures significantly more often than elongated lures. These differences may be due to such undetermined factors as texture, weight, or modifiability. The latter seems the more plausible since after the monkey selects a lure he frequently bites it, chews on it and spits it out, picks it apart manually, or merely inspects it visually.

Although the factors involved in lure-selection are not completely understood, the consistency of the preferences of monkeys seems quite similar to the consistency of preferences of monkeys for different kinds of foods and different visually explored environments.⁸ It is entirely possible that the lure has some value qualitatively similar to, but quantitatively less than, a more conventional 'reinforcer.' The third striking aspect of the findings is the lack of satiation of the response throughout practice. Although monkeys did chew the lures on occasion, the persistence of responses could hardly be due the receipt of a secondary reward. Also, the probability is quite low that each of 15 monkeys would make the same 'random' approach response on 156 consecutive occasions within a 1-min. time period.⁹

⁸R. A. Butler, Incentive conditions which influence visual exploration, *J. exp. Psychol.*, 48, 1954, 19-23; Fay, Miller, and Harlow, *op. cit.*, 13-15; H. F. Harlow and D. R. Meyer, Paired-comparisons scales for monkey rewards, *J. comp. physiol. Psychol.*, 45, 1952, 73-79.

⁹W. J. Dixon and A. M. Mood, The statistical sign test, *J. Amer. statist. Assoc.*, 41, 1946, 557-566.

EFFECTS OF MOTIVATING INSTRUCTIONS AND REPORTS OF FAILURE ON VERBAL LEARNING

By IRWIN G. SARASON and BARBARA R. SARASON, Indiana University

That variables of motivation and of learning combine to influence behavior has long been commonplace in the psychological literature. Only recently, however, have studies been made of two of the variables generally considered under the head of motivation, *i.e.* verbally induced motivation and reports of failure.¹ The present experiment was designed to determine the effect of these variables upon serial learning.

At present the effects of suggestions possessing high and low motivation are unclear. Alper studied them on immediate and delayed retention.² She found no significant differences in her immediate tests, but 24 hr. later she found that the highly motivated Ss were significantly superior to those with low motivation. Russell, on the other hand, failed to obtain statistically significant differences between groups with high and low motivation in either immediate or delayed retention.³ He did find, however, significantly greater variability of performance among the Ss with high motivation and also that reports of failure had a significant effect when retention was tested immediately after they were given but not when tested after a delay. According to him, there was no significant interaction between motivational instructions and reports of failure.

The present experiment was concerned with the effects of motivation and suggestion of failure on both immediate and delayed retention. It also permitted us to determine whether Alper's results could be reproduced with a procedure similar to hers since Russell's failure to duplicate her results may be due to his procedural modifications. Russell's Ss were tested *either* immediately *or* after a delay; Alper's Ss were tested *both* immediately and after a delay.

METHODS

Subjects. The Ss, 40 in number (22 men and 18 women), were students in introductory psychology. All were naive in respect to experiments in verbal learning. They were randomly assigned to four equal groups.

Apparatus. A memory-drum was used to expose the materials and learning was by the method of anticipation. Two lists, each of 17 nonsense syllables, were presented to S at an exposure-rate of 2 sec. per syllable with a 6-sec. interval between trials. Each list was composed of syllables randomly selected from Glaze's list of

* Received for publication August 5, 1955. This research was carried out at Indiana University. The authors express their gratitude to Dr. James A. Dinwiddie for his assistance in this study.

¹ T. G. Alper, Task orientation vs. ego-orientation in learning and retention, this JOURNAL, 39, 1946, 236-248; J. D. Lucas, The interactive effects of anxiety, failure, and intraserial duplication, *ibid.*, 65, 1952, 59-66; W. A. Russell, Retention of verbal material as a function of motivating instructions and experimental-induced failure, *J. exp. Psychol.*, 43, 1952, 207-216.

² Alper, *op. cit.*, 236-248.

³ Russell, *op. cit.*, 207-216.

75% associative value.⁴ No syllable was repeated in our lists. The score for a given trial was the number of correct anticipations.

Procedure. All the Ss from all the groups were given the following preliminary instructions.⁵

You will see nonsense syllables appearing in this opening, one at a time. After a syllable is presented, call out the next one before it appears. Of course, the first time through the list you will not be able to anticipate any of the syllables, but after the first trial call out the syllable before it appears on the screen. Prior to the beginning of each new trial there will be a short rest in which you will see blank spaces in the opening. Asterisks will indicate that the list is about to be repeated. When you see them, call out the first syllable. Do you understand?

These instructions were followed by 15 trials on the first list (List A). Then the Ss were randomly assigned to either the high or low motivational condition. The instructions for high motivation were:

Do I have your name spelled correctly? How old are you? What is your major? What class are you in? What is your grade point average? Do you expect to get a Bachelor's degree? The first list of nonsense syllables was a practice list. Here is another list. This is a short form intelligence test. It involves the memorization of syllables as in ordinary verbal learning. The list you now have to learn is, however, one which measures intelligence and ability to think in abstract terms. Pay close attention to each syllable since each one missed lowers your score when it is compared with other people of your age.

The instructions for low motivation were:

Here is another practice list. I am attempting to investigate certain characteristics of nonsense syllables, the main one being the relative associational value of the syllable. What concerns me is not your performance but rather the list characteristics uncovered. This is a common procedure in experiments with memory-drums.

Immediately after these instructions the Ss were given 14 trials on a second list of nonsense syllables (List B). The Ss in each group were then randomly assigned to one of two conditions: one, 'failure,' consisted of verbally administered instructions designed to indicate to S that he had failed; and the other, non-failure, consisted of an interval of neutral conversation between E and S about campus activities. The 'failure' instructions were as follows.

Mr. ———, you seem to be having trouble with your test. Is anything wrong? Can you see the syllables clearly? How do you feel? You've been doing worse than the other people who have worked on this task. In fact, yours is the lowest score I've gotten so far, and you are the first person I've had who has not reached the college level on this test. That's why I asked if anything was the matter. You've only gotten ——— right. Usually people get that many right in half the time it has taken you.

In the course of verbally administering the 'failure' instructions, S was shown his 'test blank' so that he could verify the number of correct anticipations that E had stated.

The interval of neutral conversation was approximately equal to the time required to administer the reports of 'failure.'

⁴ J. A. Glaze, The association value of nonsense syllables, *J. genet. Psychol.*, 35, 1928, 255-267.

⁵ The senior author was the experimenter.

After these instructions, each *S* was given an additional trial, the 15th, on List B. Then he was told to return 24 hr. later. When he reported the day following he was given five more trials on List B. At the close of the session on the second day, all *Ss* who had been given 'failure' instructions were told that their performance on the test of retention had been excellent.

RESULTS

To control individual differences among *Ss* in ability to learn nonsense syllables, analyses of covariance were made in the statistical treatment of the data. The control variable employed in these analyses was the number of correct anticipations on Trial 15 of List A.

The first analysis dealt with the effects of the differential instructions on immediate retention (Trial 15) on List B. Whereas it was found that the motivating instructions had no significant effect, the reports of 'failure' did produce a significant effect ($p < 0.025$), with 'non-failed' *Ss* giving more correct anticipations than 'failed' *Ss*. Table I presents the adjusted means for the immediate and delayed tests of retention *Ss*. The overall adjusted mean for non-failed *Ss* was 11.7 correct

TABLE I

AVERAGE NUMBER OF CORRECT ANTICIPATIONS ON TWO TESTS OF RETENTION AS AFFECTED BY MOTIVATING INSTRUCTIONS BEFORE LEARNING AND 'FAILURE' INSTRUCTIONS AFTER LEARNING

(All means corrected for level of learning in control tests.)

Motivation	Immediate test (Trial 15)		Delayed test (Trial 16)	
	failed	Non-failed	failed	non-failed
Low	7.4	12.2	5.5	6.7
High	10.8	11.2	8.4	9.8

anticipations on Trial 15 of List B, while the overall adjusted mean for failed *Ss* was 9.1. This result is in accord with the usual finding in the area of reaction to stress from failure; that is, *Ss* exposed to stressful stimuli prior to a test-trial perform at a lower level than do *Ss* who were not so exposed. The interaction between motivating instructions and reports of failure approached significance ($p < 0.06$). Failure was more detrimental for low than for highly motivated *Ss*. The adjusted means on Trial 15 of List B were 7.4 and 10.8 correct anticipations respectively for the two groups. On this same trial non-failed, low motivated *Ss* performed at a higher level (adjusted mean 12.2) than did non-failed, highly motivated *Ss* (adjusted mean 11.2).

On the delayed test for retention (Trial 16) it was found that *Ss* who had been given highly motivating instructions performed at a significantly higher level than did *Ss* given low motivating instructions ($p < 0.025$). The overall adjusted mean for high motivation *Ss* on this trial was 9.1 correct anticipations, while for low motivation *Ss* the adjusted mean was 6.1 correct anticipations. However, *F*-ratios of less than unity obtained for Failure and the Failure $\times$ Motivation interaction suggest that the effects of failure had completely dissipated during the 24-hr. interval between Trials 15 and 16 on List B.

A final analysis of covariance on the number of correct anticipations on Trial 20 of List B failed to reveal a significant effect for either Failure, Motivation, or the interaction of Failure $\times$ Motivation.⁶

DISCUSSION

That the motivational instructions and reports of failure affected the Ss' performance will come as no surprise to most workers. The specific ways however, in which the experimental treatments affected performance probably could not have been logically predicted. Certainly the 'delayed action' of the differential between motivating instructions is not readily explainable. Both the present study and Alper's experiment found that high as compared with low motivating instructions lead to superiority in performance after 24 hr., but performance does not improve immediately after administration of the instructions.⁷ This result indicates that Russell's failure to obtain a significant improvement in retention attributable to highly motivating instructions might be due to differences between his procedure and that used by Alper and in the present study.⁸ Whereas Alper, like the present experimenters, obtained immediate and delayed retention measures from all her Ss, Russell used different groups for the two measures of retention. It is interesting that such a seemingly minor procedural modification might lead to such a striking difference between Russell's results and those obtained by Alper and the present writers.

Alper has hypothesized that highly motivating instructions lead to an increase in anxiety which has a detrimental effect on the first day's performance. By the second day, Alper suggests, the inhibiting effect of anxiety dissipates and, as it were, the individual's performance reflects the beneficial effects of high motivation without the detrimental anxiety aroused on the first day. Before, however, any claim for a general superiority of high as compared with low motivation after the first day can be made, relevant parameters must be evaluated. Perhaps a more parsimonious hypothesis than Alper's might be that highly motivating instructions affect performance only to the extent to which opportunity for rehearsal takes place. This interpretation must be considered since, in the present study, the opportunity for rehearsal was greater during the 24-hr. period prior to Trial 16 than in the few minutes prior to Trial 15.

The present results concerning the effects of failure are in agreement with previous work which indicates that reports of failure result in immediate decrements in performance. The present finding that after 24 hr. this relative deficiency in performance-level for 'failed' Ss had completely disappeared would seem to support the hypothesis that reports of failure do represent a stressful situation for Ss, but that the stressfulness dissipates.

The near significant interaction between the motivational instructions and reports of failure immediately after 'failure' suggests that the effects of stressful stimuli such as 'failure' may vary depending on the motivational level of the individual. 'Failure,' in the present experiment, was more detrimental for Ss with low than

⁶ Bartlett's test failed to indicate significant differences among the variances of the different groups on Trials 15, 16, and 20.

⁷ Alper, *op. cit.*, 246.

⁸ Russell, *op. cit.*, 211-212.

high motivation, while among non-failed Ss, low motivating instructions led to better performance than did highly motivating instructions. This latter finding immediately after 'failure' stands in sharp contrast to the overall superiority of highly motivated groups 24 hrs. later.

SUMMARY

The present study was concerned with the effects of high and low motivating instructions and reports of failure or non-failure on Ss' performance in a serial learning situation. The Ss, 40 in number, were randomly divided into four equal groups: high motivation-failure; low motivation-failure; high motivation-non-failure; and low motivation-non-failure. All Ss received 15 trials on a preliminary list, immediately after which the differentially motivating instructions were administered. Following this, the Ss were given 14 trials on the experimental list and then received either reports of failure or non-failure. After these differential reports were given, Ss immediately received one more trial (Trial 15) on the experimental list. Twenty-four hours later all Ss received Trials 16 through 20 of this list.

Analyses of covariance were performed on the number of correct anticipations on Trials 15, 16, and 20 of the experimental list. The only trial for which there was a significant difference between high and low motivated groups was Trial 16, with highly motivated Ss performing at a higher level. The only trial for which there was a significant difference between failed and non-failed Ss was Trial 15, with non-failed Ss performing at the higher level.

The results indicated that the effects of the two kinds of verbal instructions (failure and motivation) on performance are markedly different. The instructions concerning failure caused an immediate but temporary decrement in performance while the motivating instructions had no immediate effect but did facilitate later performance.

ON THE REDINTEGRATIVE PERCEPTION OF WORDS

By GEORGE M. HASLERUD and R. ERNEST CLARK, University of New Hampshire

According to Solomon and Postman, "When a stimulus-pattern is presented at short durations or at low illumination-intensities, only fragments of the total word-stimulus are 'effective.' Such a stimulus-fragment may be considered to represent a point on the generalization-dimension of stimulus-patterns capable of eliciting the correct verbal response."¹ Certain questions regarding this statement immediately present themselves: (1) Where in the word are the stimulus-fragments that are the effective cues? (2) What conditions attend the redintegrative leap from the fragment to the correct whole-response?

A partial answer to the first question is given by Erdmann and Dodge, Wagner, and others who found in nonsense arrangements of letters that the first and last letters were often correctly recognized while the middle letters were not.² This finding needs to be checked for meaningful words. Also only with meaningful words can the second question be answered. Words of length greater than the individual's span of apprehension for letters were chosen as the exposure-material and were so chosen that the responses ranged from correct to fragmentary reports.

Ever since Cattell's experiments, the crux of the problem seems to be why familiarity and meaning make it possible for a *S* to surpass his span for unconnected letters.³ One may experimentally induce such superiority by 'building in' familiarity, as in Solomon and Postman's experiment, or by using extraneous organization, as in Vanderplas' study.⁴ A direct, immediately available measure of familiarity can, however, be found in the individual's ability to define a word. Are words, as equated in Thorndike's frequency of usage, perceived more readily in short exposures when *S* can define them than when he can not? Does the ability to define have any relation to the ability to recognize the stimulus-word among alternatives presented after an unsuccessful recognition? The study of the function of various aids to recognition may provide a solution to Cattell's problem.

Procedure. The *Ss* were 25 students (12 men and 13 women), freshmen in liberal arts and agriculture.

The stimulus-materials for the tachistoscopic experiments were 22 nine-letter words which were selected as follows: two words were chosen from the Thorndike list from the tenth to the twentieth thousand;⁵ words beginning with j, k, q, x,

* Received for publication January 18, 1956.

¹ R. L. Solomon and Leo Postman, Frequency of usage as a determinant of recognition threshold for words, *J. exp. Psychol.*, 43, 1952, 199.

² R. S. Woodworth, *Experimental Psychology*, 1938, 727-745.

³ *Ibid.*, 738.

⁴ J. M. Vanderplas, Frequency of experience versus organization as determinants of visual thresholds, this JOURNAL, 66, 1953, 574-583.

⁵ E. L. Thorndike, *A Teachers' Word Book of the Twenty Thousand Words Found Most Frequently in General Reading for Children and Young People* (Rev.), 1932.

y, and z were, however, rejected. The words selected and their Thorndike levels were: 10,000, nightmare, unstained; 11,000, recessive, ascension; 12,000, collusion, whirligig; 13,000, secession, economize; 14,000, fecundity, pecuniary; 15,000, transmute, barrister; 16,000, unsullied, declivity; 17,000, microcosm, guileless; 18,000, nefarious, itinerant; 19,000, hatchment, obtrusive; 20,000, lorgnette, virulence. The words were printed in capitals at the center of white cards (3×4 in.) and in a type that brought them to about 2 in. in length.

The cards were exposed singly 15 in. from the nosepiece of a photographic shutter-type tachistoscope set to allow its 50-w. lamp to light the stimulus-card with 0.2 candles of incident light per sq. ft. at the nosepiece. It is assumed that the 40-m. sec. exposure we used would restrict *S* to a single fixation and would so impoverish the stimulus that there would be fragmentation. The average threshold for nine-letter words of difficulty similar to our list has been found to be 96 m. sec.⁶

The material for the multiple-choice recognition-test consisted of the stimulus-word and four others that began and ended with the same letters as the stimulus-word. The other words were from approximately the same thousand of the Thorndike list such as: mannerism, mesmerism, multiform, mysticism, microcosm. The words were typed on one side of 5×7 in. cards, and on the reverse side of the card was the stimulus-word with four possible definitions, one taken from Webster's Collegiate Dictionary, and three decoy definitions. For example: Microcosm—a magnifying device, a philosophy of religion, the solar system, compact representation. These multiple-choice definitions undoubtedly gave more positive identifications than would have unaided recall, but since each *S* was his own control, the relative difference between the defined and non-defined words was adequate for our purpose.

A practice session on seven-letter words in the tachistoscope permitted the *S* to learn how to keep his eyes light adapted, to fixate at the center of the stimulus-location and, immediately after the exposure, to report the word or any part of it. In the experiment proper the testing sequence for each nine-letter word was: a 40-m. sec. exposure of the printed word; immediate report of what was seen (if no word, the letters and their position); presentation of the multiple-choice recognition-test with instruction to pick the stimulus-word from the five alternatives; and then turning over the card to have the *S* pick from four alternative definitions the one that was correct for the stimulus-word typed at the top of the card.

To test all *Ss* with words of approximately equal difficulty, all *Ss* were started on the pair of words at the ten-thousand level. If these were successfully defined, the next sampling was at the thirteen-thousand level, and, if these were defined, the sixteen-thousand, and so on. After the first mistake in definition, the adjacent words up and down the levels of difficulty were presented until eight correct definitions and eight mistaken definitions had been secured. When more than 16 words had to be presented, those closest to the first mistake were used. The results on these 16 words from the tachistoscopic and multiple-choice tests make up the data for each *S*. The absolute frequency of usage of the word is irrelevant to our problem. Words defined and not defined overlapped considerably in frequency on the Thorndike list; the mean of the lower ranges of the defined words, 12.44

⁶ Elliott McGinnies, P. B. Comer, and O. L. Lacey, Visual-recognition thresholds as a function of word length and word frequency, *J. exp. Psychol.*, 44, 1952, 65-69.

thousands, is almost identical to that for the non-defined words, 12.68. At the other end of the range the means (18.28 and 19.48) are significantly different.

Results. The results, given in Table I, show that over three times as many defined words were correctly perceived in immediate unaided recognition as were words that could not be defined. This trend was the same for Ss of both sexes. The men and women differed, however, in the number of words correctly perceived. The women's mean number of words recognized was 6.46 out of 16; the men's was 4.50; the difference was significant at the 2% level.

TABLE I
NUMBER OF STIMULUS-WORDS CORRECTLY PERCEIVED OUT OF EIGHT

Test	Group	N	Condition	Mean	SD*	Diff.	SE _{diff} †	t	p
Unaided recognition	M	12	Definition Known	3.67	1.71	2.84	.51	5.55	<.001
			Not Known	.83	.83				
	F	13	Definition Known	4.92	1.20	3.38	.47	7.19	<.001
			Not Known	1.54	1.38				
Recognition aided by alternatives	M	12	Definition Known	7.08	1.10	.16	.40	0.4	>.1
			Not Known	6.92	1.05				
	F	13	Definition Known	7.23	.84	.08	.27	0.3	>.1
			Not Known	7.15	.93				

* Corrected for small samples.

† By difference method for correlated pairs.

In contrast to the results on the tachistoscopic test of recognition were those on 'aided recognition.' Generally, the stimulus-word was chosen 85-90% of the time correctly from among the five alternatives, but, as shown by a *p* greater than 0.1, the difference between words that could and could not be defined is now insignificant.

The number of correct letters in each of the nine positions in the words was analyzed for the 400 responses to the tachistoscopic presentation, as shown in Fig. 1. The responses were classified into four groups: correctly perceived stimulus-word, 34% of the total; failure to recognize the stimulus-word but response a somewhat meaningful word, 41%; fragmentary letters or groups of letters, 21%; and finally, no correct letters, 4%. The curves for the percentage of correct letters in the incorrect words and letter fragments are all U-shaped for both sexes. The minimal number correct for the first letter was 81% and for the ninth letter, 75%, but the middle of the curve sags to a mere 5% when only fragments can be reported. For the incorrect words 40% to 60% of the middle letters are correct. Apparently the fragmental cues most likely to be effective are at the ends of the word despite fixation at the middle. Moreover, they are apt to be at both ends; out of 84 responses of only fragments, 51 were two-ended (significant at the 5% level).

Discussion. That the beginning and ending letters of the word are perceived correctly even when nothing else can be reported correctly is obviously a function of their physical position. Perhaps the significant factor is the sharp gradient between

the word and its white ground.⁷ If the word were embedded in a longer line of similar letters, this stimulus-gradient would not be available to make the end letters supraliminal. The words would have only such aids to recognition as are available to the middle letters of the present experiment.

In the multiple-choice test of our experiment, the end letters were cancelled out as a differential cue by making them all identical. Nevertheless, the stimulus-word was chosen correctly in 85-90% of the trials, indicating that under these conditions the previously masked middle letters may aid the perception. Partly, it may be that the meaning of the word functions as an aid in reporting. It is difficult to

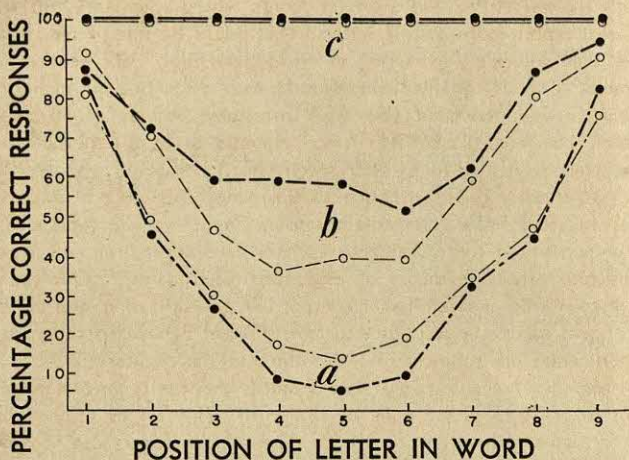


FIG. 1. PERCENTAGE OF LETTERS REPORTED CORRECTLY FOR VARIOUS TYPES OF RESPONSES

The open circles are for women, the filled are for men. *a* = fragmentary responses; *b* = incorrect whole; and *c* = correct whole responses.

recall without confusion long strings of unrelated letters. The knowledge of the word may serve to reduce the stimulus to one of a small number of possible items, and less information will be required to perceive it correctly than when it is one out of a larger number of possible items. The reduction in the number of alternatives can apparently be equally effective before and after the stimulation.⁸

Now if familiarity is thought of as the result of learning "to identify the stimulus relative to all possible stimuli,"⁹ the aided recognition-test may be a prototype of the condition for any correct perceptual response. This test provides an explicit set of alternatives as 'possible stimuli,' but previous learning of familiar words will

⁷ J. J. Gibson, *The Perception of the Visual World*, 1950, 112-114.

⁸ D. H. Lawrence and G. R. Coles, Accuracy of recognition with alternatives before and after the stimulus, *J. exp. Psychol.*, 47, 1954, 208-214. For an analysis of the basis of tachistoscopic reports see also A. D. Glanville and K. M. Dallenbach, The range of attention, this JOURNAL, 41, 1929, 207-236, esp. 233-236.

⁹ J. J. Gibson and E. J. Gibson, What is learned in perceptual learning?, A reply to Professor Postman, *Psychol. Rev.*, 62, 1955, 450.

also limit quite as much the possible combinations of letters. Support for such a view is found in the 41% of incorrect but somewhat meaningful word-responses. Although the number of correct end-letters is approximately the same as for the fragmentary responses, the middle of the meaningful words contains significantly more correct letters. Our curves for this type of response are very much like those of Wagner who, although he used a nonsense combination of eight letters, emphasized in his instructions to his subjects the wholeness of the combination.¹⁰ Probably through limiting the number of possible alternatives the very attempt to give a word as a response aids the perception of the middle letters.

The curves for the sexes are much alike when the responses are fragments, contrasting with the curves for the incorrect words where a real sex difference may exist. The men report more middle letters. This might be interpreted, however, as evidence of the women's superiority in word perception. The tendency to give word responses is about equally strong for the two sexes, but the men's responses include many combinations that just miss correctness because of mistakes in the middle letters. The women gave 84 correct responses to 53 for the men, while the ratio of incorrect words is 70 to 92 respectively.

Postman has recently called attention to the plausibility of a neural redintegrative process initiated by a fragmentary sensory input, yet he concluded, "However, the components of the redintegrative chain have never been sufficiently specified to constitute a testable theory of mediating mechanisms."¹¹ Our experimental results point away from a single factor concept like generalization, and instead, show the need for two successive responses to two different behavioral supports, the one in the experimental environment and the other in the organism.

With words that are subliminal unless aided, the environmental factor is the abrupt contrast-gradient at the ends of the word. This might offer sufficient information for later elaboration by the S. The middle letters can, however, also provide cues and become supraliminal if aided by the second factor. This is the restricted set of alternatives provided either in the experimental situation, or in the verbal repertoire of the S. We found evidence for this in our aided recognition where the beginning and ending letters were controlled and only the middle of the word could provide differential cues.

¹⁰ Julius Wagner, *Experimentelle Beiträge Zur Psychologie des Lesens*, *Z. Psychol.*, 80, 1918, 1-75. Summarized in Woodworth, *op. cit.*, 742-743.

¹¹ Leo Postman, Association theory and perceptual learning, *Psychol. Rev.*, 62, 1955, 443.

A NEW EXPLANATION OF THE VELOCITY-TRANSPOSITION PHENOMENON

By OLIN W. SMITH and LOUISE SHERLOCK, Cornell University

J. F. Brown discovered a velocity-transposition phenomenon which he defined as follows: "If a moving field in a homogeneous surrounding field is transposed in its linear dimensions as 1:2, the stimulus-velocity must be transposed by a like amount in order that the phenomenal velocity in both cases be identical."¹ Brown's statement implies that factors in addition to the physical velocity of the moving objects determine phenomenal velocity. Brown investigated some of these factors and others have been more recently discussed by Cartwright, Wallach, and Smith.²

One hypothesis which has been overlooked previously is that the judgments of apparent velocity were determined by the frequency at which the moving objects left their bounded fields. Movement is a necessary condition for such a frequency, but, when physical velocity is held constant, this type of frequency can be varied independently by varying the density of discriminable objects in the moving field. First of all, if this type of frequency were the stimulus-correlate for judgments of apparent velocity, then judgments of the frequency of objects passing a fixed point in Brown's situation should yield data as predicted by the transposition-principle. Secondly, if the judgment of the frequency of objects passing a fixed point can be made without making any judgment of physical velocity, then instructions to set the variable at an appropriate ratio of frequency should result in a physical velocity of the variable which closely approximates the physical velocity of the standard.

The purpose of Part I of this study is to investigate the above two predictions. If judgments of frequency of objects passing a fixed point yield transposition data, then the hypothesis is tenable that the stimulus-correlate of apparent velocity as investigated in Brown's experiment is apparent frequency. Brown's transposition phenomenon could be then explained on the basis of judgments of frequency, a simpler explanation and more general in that there would be no necessity for invoking transposition as one of the necessary conditions of the judgment.

When size-judgments are made of test-objects seen in completely homogeneous surrounds, retinal rather than veridical size-matches have been repeatedly obtained.³ From this would be predicted that motions similarly observed and judged would yield matches of retinal velocity rather than physical matches. The purpose of Part

* Received for publication October 14, 1955. This paper was supported in part by Contract Nonr401(14) with the Office of Naval Research.

¹ J. F. Brown, The visual perception of velocity, *Psychol. Forsch.*, 14, 1931, 199-232.

² Dorwin Cartwright, On visual speed, *Psychol. Forsch.*, 22, 1938, 320-342. Hans Wallach, On constancy of visual speed, *Psychol. Rev.*, 46, 1939, 541-552. O. W. Smith, The effects of windows of two sizes on matches of objective velocity (in press).

³ A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance variant, this JOURNAL, 54, 1941, 21-37; A. H. Hastorf and K. E. Way, Apparent size with and without distance cues, *J. gen. Psychol.*, 47, 1952, 181-188; William Lichten and Susan Lurie, A new technique for the study of perceived size, this JOURNAL, 63, 1950, 280-282.

II is to test the prediction that retinal matches of velocity would be made when the movement fields are observed in the absence of any cues to distance, external to the fields.

Apparatus. Two devices were used to produce motion.⁴ In each an endless rubber belt moved through a flat display area. The velocity of either belt was continuously variable by remote control of either *E* or *S*. Each belt was covered by white paper. Black, plastic tape, 1 in. wide, made alternate black and white horizontal stripes across the surface of one belt. The second belt was similarly covered by $\frac{1}{2}$ -in. black and white stripes. Each was illuminated by a pair of 40-w. cool light, fluorescent tubes. In front of each was a clear plate glass window to which was stuck a very small piece of the black tape.

At *S*'s position were a table and chair completely shielded by a black cloth from the devices producing motion. A rotary control crank by which *S* varied the velocity of the $\frac{1}{2}$ -in. striped surface was mounted on the table. The eyepieces of two Imarect finders protruded through the cloth. By moving his head slightly *S* could view either surface. The finders contained variable rectangular apertures. The 1-in. and $\frac{1}{2}$ -in. striped surfaces measured approximately 23 in. x 17 in. and 11.5 in. x 8.5 in. respectively, and both were 6 ft. from the finders. The spots of black tape on the two windows appeared to be in the center of the two areas. The physical conditions of the experiment on transposition were completely satisfied. Both fields were in homogeneous surrounding fields (a blackout) and one was transposed in all its linear dimensions as 1:2.

Experimental plan. Each of 24 *Ss*, who were paid for their services, made three types of judgments. Judgments A and B were returned in Part I of the study and C in Part II. In Judgments A, *S* matched the frequency at which the 1-in. lines passed behind the fixed center point by adjusting the velocity of the belt bearing the $\frac{1}{2}$ -in. lines. The physical velocity of the standard, the 1-in. lines, was constant at 20 ft. per min. for all three judgments. A correct match for equal frequency on the variable was a velocity of 10 ft. per min., thus fulfilling Brown's specifications for identical phenomenal velocities in the transposition-situation. In Judgments B, *S* adjusted the frequency of the $\frac{1}{2}$ -in. lines to twice the frequency of the 1-in. lines. A correct match on the variable would be 20 ft. per min. In Judgments C, *S* was instructed to regard the smaller lines as being twice the distance of the larger lines. They were to take this into account in making matches of physical velocity. Four orders of judgments were used: ABC, BAC, CAB, and CBA. Six *Ss* were assigned at random to each order. Judgments C were made 12 times first to permit a comparison for order. No *S* was informed of the correctness of his judgments.

Instructions. The instructions, the order of which was adjusted to suit the order of judgments assigned the *Ss*, were as follows:

For Judgments A: In front of you are two apertures. Use your right eye in looking in each of these. Look in the aperture on your left. You will see a surface on which the black parallel lines are in motion from bottom to top. You will also

⁴O. W. Smith and J. J. Gibson, Apparatus for the systematic variation of visual translatory motion (in press).

see a little black point which is stationary in the middle of your field of view; looking through the aperture on the right you will see another surface with smaller parallel black lines in motion from bottom to top and a stationary black point in the center of your field of view. By rotating the speed control crank you can slow down or speed up the frequency at which the lines on the surface in front of you go past the stationary black point. The lines on the surface to your left pass the fixed point at a constant frequency. Adjust the apparent frequency of the lines passing the stationary black point on the surface to your right so that it equals the apparent frequency of the lines passing the stationary black point on the surface to your left. Do not count. We wish you to make the match purely on the basis of the apparent frequency at which the lines go by the point.

Before each judgment by each *S*, *E* set the physical velocity of the variable either far too high or barely moving.

For Judgments B: The instructions are the same as for Judgments A except now your task is to adjust the apparent frequency of the lines passing the stationary black point in the center of the surface to the right so that it is twice the frequency of the lines passing the stationary black point on the surface to your left.

If Judgments B were called for first, the instructions for Judgments A were read with the exception of the section concerning *S*'s task.

For Judgments C: There are two apertures in front of you. Use your right eye when looking at either one. By looking through the aperture to your right you will see a surface in motion from top to bottom. The same is true if you look through the aperture on your left. The speed of the moving surface on your left is constant. By using the speed control crank you can slow down or speed up the velocity of the surface on the right. In that aperture you will note that the moving surface appears to be twice as far away from you as the surface on your left. Take this factor into account and adjust the velocity of the surface in front of you so that it is equal to the physical velocity of the surface on your left. Do not count, or make use of time judgments for your estimates. You are to equate the physical velocities of the two surfaces.

Results and discussion: Part I. The means and standard deviations in feet per minute (ft./min.) of the measured settings of the variable surface for Instructions A and B were 10.52 ± 2.52 , 21.27 ± 6.02 ft./min., respectively. The mean of the matches of equal frequency, Instruction A, was approximately the value predicted by Brown for phenomenal identity of velocity. The mean of 10.52 ft./min. was not significantly different by *t*-test from 10 ft./min., an exact match of frequency. The mean of 21.27 ft./min. was not significantly different by *t*-test from 20 ft./min., the velocity of the variable for a frequency twice that of the standard.

Part II. The mean and standard deviation in feet per minute of the measured settings of the variable for Instruction C for matches of physical velocity were 9.65 ± 2.18 respectively. No significant order-effects were found, so the 24 judgments were pooled. The mean of 9.65 ft./min. was significantly different by *t*-test from 20 ft./min., a retinal angular velocity match. It was *not* significantly different from 10.52 ft./min., the mean match for equal frequencies. When Brown had his *Ss* judge the apparent velocity of a 'transposed' stimulus-pattern in which one field was at twice the distance from the *S* as a second field, he obtained data illustrating 'transposition.' The simpler hypothesis, independent of transposition, is that *Ss* made frequency-matches when instructed to make matches of either apparent or physical velocity under these conditions. The frequency hypothesis is preferred,

since frequency can be an independent variable when retinal velocity is held constant.

Conclusions: Part I. (1) The hypothesis is tenable that when Brown's Ss were instructed to make judgments of apparent velocity, they were making matches of apparent frequency. (2) Frequency can be judged independently of physical velocity.

Part II. (1) When the 'cues' to the distances of two moving fields are completely eliminated, 'retinal' matches were not made as predicted by data of judgments of size under similar conditions. (2) An explanation dependent upon frequency-judgments was preferred since they can be related to a variable of stimulation when retinal velocity is held constant.

APPARENT VISUAL SIZE AS A FUNCTION OF DISTANCE FOR CHILDREN AND ADULTS

By H. PHILIP ZEIGLER and H. LEIBOWITZ, University of Wisconsin

There is some question in the literature as to whether size-constancy varies with age.¹ Data are, however, lacking from an investigation of size-constancy for which a wide range of distances is available and for which a functional relationship can be plotted. The present study was designed to obtain such results.

Apparatus and procedure. The procedure was modeled after that used by Holway and Boring.² A series of standard stimulus-objects, made from 1-in. diameter wooden dowels, were prepared such that at the distances used in the study, the objects subtended a visual angle of 0.96° at *S*'s eye. The comparison-object, also a 1-in. diameter dowel, was so arranged that the visible portion of its length could be continuously varied by moving it up or down through a hole cut in the center of a board. The remainder of the comparison-object was hidden from *S* by a suitably placed curtain. *S* was seated at a distance of 5 ft. from the comparison-object, which was approximately at eye-level. The standard objects were fitted into small black wooden blocks and mounted on a black metal stand 3 ft. from the floor in front of and slightly to one side of the comparison-object.

The experiment was conducted in a room 108 x 22 ft. From *S*'s position three windows were visible on one side of the room and one at the extreme end. There were several pieces of furniture visible along the walls. Illumination was provided by six 100-w. bulbs in addition to the windows, and the experiments were conducted in the early afternoon and only on sunny days.

Five distances were used; 10, 30, 60, 80, and 100 ft. For each distance *S* made four judgments, two ascending and two descending. (The actual manipulation of the comparison stimulus was done by *E* at *S*'s direction.) The distances were presented in a random order and, between successive trials with new distances, *S* was so blindfolded that he could not see *E* setting up the standard objects.

The instructions for both children and adults were as follows.

I am going to move this stick (pointing to the comparison-object) up and down. I want you to tell me when it looks as high as the one out there (pointing to the standard object).

S was told to disregard the stand or the wooden block and to base his judgments on the height of the standards. *E* avoided giving any information as to whether the

* Received for publication November 14, 1955. Supported in part by Research Grant M-1090, National Institute of Mental Health. The authors thank the staff of Lenox Hill Neighborhood House, New York City, for their cooperation in obtaining the *Ss* used in this experiment.

¹ Franz Beyrl, Ueber die Grössenauffassung bei Kindern, *Z. Psychol.*, 100, 1926, 344-371; H. Frank, Untersuchung über Sehgrössenkonstanz bei Kindern, *Psychol. Forsch.*, 7, 1926, 137-145; Die Sehgrössenkonstanz bei Kindern, *ibid.*, 10, 1928, 102-106; Wilhelm Burzlaff, Methodologische Beiträge zum Problem der Farbenkonstanz, *Z. Psychol.*, 119, 1931, 177-235.

² A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance variant, this JOURNAL, 54, 1941, 21-37.

standards were of the same or different sizes. S was also instructed to make all judgments with binocular regard and to wear glasses if needed.

The Ss were 13 in number (8 boys, ranging from 7 to 9 yr. of age, and 5 men, from 18 to 24 yr. of age) all resident at a summer camp. The men were told they were to serve as controls for the experiment with the children. For the boys, the experiment was part of a 'scouting' game which included outdoor as well as indoor estimates of the sizes of various objects.

Results. The results for the children are given in Table I and for the adults in Table II. Both sets of data are plotted in Fig. 1. On this plot a horizontal line

TABLE I
MATCHED SIZE (IN INCHES) OF STANDARD STIMULUS-OBJECT AS A FUNCTION OF DISTANCE
(Children)

S	Age (in yr.)	Distance of standard (in feet)				
		10	30	60	80	100
		Size of test object (in inches)				
		2	6	12	16	20
1	9	1.3	3.3	2.9	3.5	3.3
2	8	1.9	4.3	7.8	8.2	9.2
3	8	1.5	3.1	5.3	6.4	5.5
4	7	2.9	4.1	6.6	7.8	9.4
5	8	3.1	5.1	9.9	11.1	13.1
6	8	2.5	5.2	8.3	8.0	8.3
7	8	2.1	4.8	11.1	8.2	10.9
8	9	.8	1.1	1.4	1.2	1.1
Mean		2.0	3.9	6.7	6.8	7.6
SD		.8	1.3	3.3	3.6	4.0
Brunswik ratio		1.00	.58	.52	.39	.35

represents the law of the visual angle; a theoretical condition in which perceived size can be predicted on the basis of geometrical optics. The law of size-constancy is represented by a line of slope 0.017 ($\tan 0.96^\circ$). If perceived size were independent of distance, *i.e.* perfect size-constancy, the data would fall along this theoretical function.

It can be seen from Fig. 1 that the data for the adults fall very close to the prediction in terms of size-constancy as previously reported by Holway and Boring. The data for the children however, fall at positions on the matched size-axis closer to the line representing the law of the visual angle. Differences between the means of matched sizes for adults and children were calculated for all distances above 10 ft. A Mann-Whitney non-parametric u -test indicates that these differences are significant ($u < 0.01$ for 30, 80, 100 ft.; $u < 0.05$ for 60 ft.).

Discussion. The results of the present investigation demonstrate closer agreement with the law of size-constancy for a group of adults than for a group of children. Such results are consistent with the conclusions of Beyrl, but differ with respect to the magnitude of the difference in size-constancy.³ As has been pointed out by

³ Beyrl, *op. cit.*, 369.

Frank, all of Beyrl's Ss, including his 2-yr. old children, demonstrate rather high constancy.⁴ Brunswik ratios computed for his data, which were obtained at distances between 6 and 37 ft., for the 2-yr. old group range from 0.77 to 0.92 for a disk test-object, and from 0.87 to 0.95 for blocks. For his 8-yr. old group ($n=2$) the ratios vary from 0.97 to 0.99 for both kinds of test-objects. Brunswik ratios for the children in the present study are 1.00 at 10 ft., 0.58 at 30 ft., and drop to 0.35

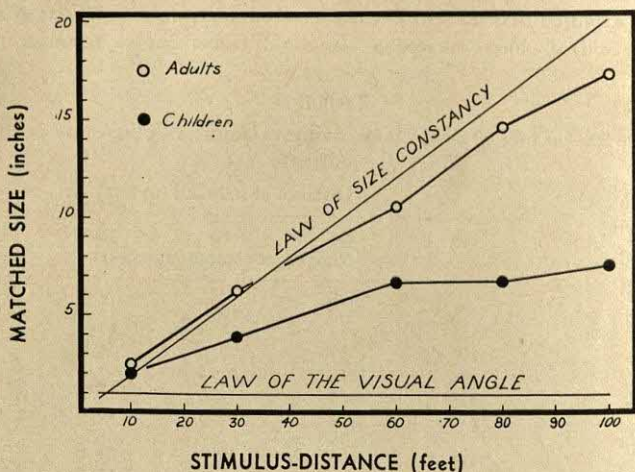


FIG. 1. MEAN MATCHED SIZE AS A FUNCTION OF STIMULUS-DISTANCE FOR A GROUP OF ADULTS AND OF CHILDREN

The Size of the Test-Object Was Adjusted so as Always to Subtend a Constant Visual Angle.

at 100 ft. Thus, the present study agrees with Beyrl's conclusion that size-constancy is not as well developed in children as in adults, but differs in that for comparable stimulus-distances the children show less constancy than did Beyrl's.

In addition, the present experiment, utilizing a greater range of distances, indicates that the differences between children and adults increases with the distance of the test-object. This can be clearly seen from the shape of the curve in Fig. 1. The function for the adults is nearly linear, and lies close to the line representing the law of size-constancy. The curve for the children rises less rapidly with distance and approaches a limiting value at about 60 ft. The differences between the mean matched size at 60 and 100 ft. are not significant for the children ($p>0.05$) and are significant for the adults ($p<0.01$).

With respect to inter-S variability, the present results are consistent with those of previous investigators. As has been pointed out by Smith, studies of size-constancy typically demonstrate an increase in variability with distance.⁵ That this also holds

⁴ Frank, *op. cit.*, 1928, 103.

⁵ W. M. Smith, A methodological study of size-distance perception, *J. Psychol.*, 35, 1953, 143-153.

true for the present investigation is evident from inspection of the standard deviations in Tables I and II.

The present results, obtained over a range of 90 ft., demonstrate that the functional relationship between matched size and distance is different for adults and children. Such results are interpreted as supporting the view that size-constancy increases with age.

Summary. The function relating matched size to distance was determined for groups of children and adults. The Ss matched a comparison-object to one of a series of standard objects located at various distances but so adjusted in size as

TABLE II
MATCHED SIZE (IN INCHES) OF STANDARD STIMULUS-OBJECT AS A FUNCTION OF DISTANCE
(Adults)
Distance of standard (in feet)

S	Age (in yr.)	Size of test object (in inches)				
		10	30	60	80	100
		2	6	12	16	20
1	20	2.2	6.4	8.9	15.1	15.7
2	24	2.6	6.2	7.9	9.8	15.0
3	18	2.4	5.9	10.4	11.6	12.3
4	20	2.3	5.6	13.6	18.8	21.1
5	19	2.9	7.4	12.4	17.8	22.6
Mean		2.5	6.3	10.6	14.6	17.3
SD		.36	.78	2.3	3.8	4.3
Brunswik ratio		1.50	1.06	.87	.91	.86

always to subtend a constant visual angle. The results for the adults lie close to the theoretical line representing the law of size-constancy. The results for the children were more in agreement with the law of the visual angle. The adult-curve is nearly a straight line, while that for the children approaches a limiting value at 60 ft. These data are viewed as lending support to the thesis that size-constancy increases as a function of age.

THE PREDICTION OF COLLISION

By ROBERT M. GOTTSDANKER and RALPH V. EDWARDS, Santa Barbara College

No direct information is available on the factors which enter into an operator's decision as to whether two moving objects will collide, nor are there data on the accuracy with which such predictions are made. The previous studies which bear the closest relation to these problems have involved only one moving object. In a motion-picture *Estimation of Velocity Test* devised by Gibson, S was required to estimate at a subsequent instant of time the position of a moving airplane which had disappeared behind a cloud.¹ The responses indicated that S generally was able to extrapolate the target-motion with an error of velocity of less than 20%. Experimental work by one of the present writers, using a task in which S was required to continue tracking after the target disappeared, found the average error of continuative velocity to be from 12 to 17% of the previous velocity.² The continuation of accelerated patterns of motion also was studied. It was found that the continuation maintained a constant velocity which was lower than the last visible velocity for positively accelerated targets and higher for negatively accelerated ones. It was inferred that S had either used a particular 'stale' velocity or else had integrated velocities over some period of time.

In the present experiment, one type of collision-situation was studied. Two targets in marked orthogonal lanes approached an intersection, and a centrally-placed S predicted their future relative positions. There are, of course, many important collision-situations to which one or more of these specifications will not apply, and it should be emphasized that each variation in situation may well bring new factors into the judgments. Expectations here were based upon the earlier studies on estimation and continuation. It was predicted that judgments of the relative future positions of two moving targets would reflect an error of 15 or 20% in estimating the velocities. The smoothing of accelerated motion also was foreseen.

METHOD

Subjects. The Ss were 10 men between the ages of 17 and 35 yr. Six of the group were college students.

Apparatus. The apparatus was a modification of the paper-and-pencil tracking box, described previously, in which sheets of paper with printed target-lines are driven past a viewing slot in a concealing mask.³ The major change from the

* Accepted for publication February 6, 1956. This research was supported by the U. S. Air Force under contract AF 33(616)-2024 with the University of California monitored by the Psychology Branch of the Aero Medical Laboratory, Directorate of Research, Wright Air Development Center. Permission is granted for reproduction, publication, use, and disposal in whole or in part by or for the U. S. Government.

¹J. J. Gibson (ed.), *Motion Picture Testing and Research*, U. S. Army Air Forces Aviation Psychology Program, Research Report No. 7, 1946, 62-65.

²R. M. Gottsdanker, The accuracy of prediction motion, *J. exp. Psychol.*, 43, 1952, 26-36; Prediction-motion with and without vision, this JOURNAL, 65, 1952, 533-543; A further study of prediction motion, *ibid.*, 68, 1955, 432-437.

³Gottsdanker, *op. cit.*, *J. exp. Psychol.*, 43, 1952, 26-36.

earlier arrangement was that instead of using the single viewing slot in the concealing mask, set perpendicular to the direction of motion of the paper, two slots were used. They formed a 90° intersection, each slot making a 45° angle with the direction of the motion of the paper which passed underneath. In this manner, when two lines were printed in the proper position on the paper, each generated a moving target in one of the slots. Of course, a line could not be allowed to extend inward beyond the center point, since that would make a considerable segment of the line visible at once in the wrong slot. The paper moved at a constant velocity of 22 mm. per sec.

Patterns. In all, eight different stimulus-patterns were used. Each consisted of a standard line and a variable line sloping to a stippled 'cloud.' On four of the patterns the variable line was to S's left, and generated a target in the slot

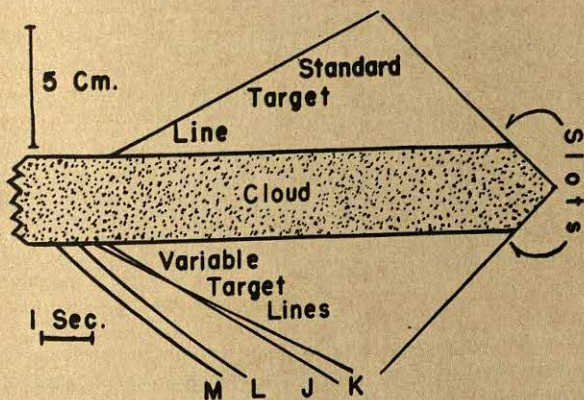


FIG. 1. STANDARD AND VARIABLE PATTERNS

surrounded by a black outline. On the other four patterns the variable line was to S's right and generated the target in the white slot. The standard was always on the side opposite the variable. Fig. 1 is a scale drawing of the four patterns. S stood before the apex of the lines marked 'Slots.' Each run was started with the target-sheet so placed that these lines appeared in the slots of the mask. Actually, only one variable line was paired with the standard on each pattern. In all treatments of the data, the responses made to each of the four patterns indicated were combined with those made to its mirror-image. Future designations will be in terms of the four combined patterns rather than the eight single-sided ones.

On all patterns, the target generated by the standard line moved at a constant speed of 15.5 mm. per sec. Two of the four variable targets (J and K) also were linear. As shown in Fig. 1, they appeared in the viewing slot shortly after the standard but had slightly higher velocities. The other two target-lines were portions of parabolas which started more slowly than the standard but finished more rapidly. They were of identical form but slightly displaced in time. They appeared in the viewing slot later than the standard. It may be seen in Fig. 1 that the targets were visible until they disappeared into the 'cloud.' Had the standard target been

allowed to continue at its previous velocity, it would have reached the point of intersection 2.8 sec. after it disappeared.

Task. *S* was read the following instructions:

Your task is to judge where the line in the white slot will be at the exact instant the line in the black slot reaches the point of intersection. As soon as you are able to make a judgment, mark your estimate with your pencil. Keep the pencil ready to use and, once you start marking, hold the pencil stationary on the paper.

Measurements. Each of the runs was scored by measuring the difference between the red line marked by *S* and a constructed center line running the length of the cloud. This value was easily converted into distance from intersection in the slot by multiplying by $\sqrt{2}$. It should be emphasized that *S* had considerable distance beyond as well as before the point of intersection in which he could make his mark. A choice beyond this point was given a *plus* sign, indicating that *S* judged that the variable target would already have crossed the intersection by the time the standard target arrived at that point. The *minus* sign designated a choice before the intersection.

As the placement of the target-sheet in relation to the point of intersection was critical in generating the proper display, registration-marks were made on each run to determine whether the motion was within acceptable limits. Unfortunately, it was necessary to eliminate 269 of the total of 800 runs.

Sequence of trials. Each of the eight single-sided patterns was presented to *S* once as a practice-trial. Following the practice trials, each pattern was presented 10 times in random order. The entire procedure required about 40 min. for each *S*.

RESULTS

Discriminations among patterns. For every *S*, a mean and standard deviation of predicted positions was computed for the acceptable runs on each of the four combined patterns. Next, still at the level of the individual *S*, the six comparisons in mean predicted position among the four patterns were subjected to the standard *t*-test for independent groups. There was considerable difference among *S*s in the extent to which they made differential responses to the four patterns (5% level). *S*₁ and *S*₆ made differential responses in five of the six comparisons, while at the other extreme, *S*₂, *S*₃, *S*₄, and *S*₁₀ made no differential responses. It should be noted further that 21 of the 24 discriminations represent judgments of farther motion relative to the standard of patterns *J* or *K* as compared with *L* or *M*.

Analysis of responses of discriminating Ss. The over-all mean predictions of the six *S*s making some differential responses were -3.0, -4.0, -10.3, and -14.0 mm. for the *J*, *K*, *L*, and *M* patterns respectively. These negative values indicate that all mean predictions were that the variable target would not yet have reached the intersection when the standard target arrived. Evidently the predictions do not at all correspond to the relative positions which would have resulted from objective extrapolation of terminal velocities, which were +2.5, -1.8, +3.0, and -2.8 mm. for the patterns in the same order. Rather, they are in very close agreement with their position relative to the standard at the time of reaching the cloud: -1.6, -3.0, -9.5, and -15.2 mm.

Precision of response. Mean values of the *SD*s for the 10 *S*s are 4.49, 4.12, 6.93, and 6.74 mm. for patterns *J*, *K*, *L*, and *M*, respectively. When these values

are converted to percentages of the velocity of the standard target, the relative indices of precision are 7.3, 6.7, 11.3, and 10.9%.

Discussion. Expectations that differential velocities would be utilized in predictions, and that accelerated motion would be smoothed, were not realized in the present findings. For both accelerated and constant-velocity targets, prediction was based instead on relative positions at the time of target-disappearance. One may conjecture that the consequence of increased complexity of the predictive situation is to decrease the order of the predictive equation employed by *S*. No great importance can be attached to indices of the precision of choice which are qualitatively inadequate (as predictions based upon relative position must be for targets which differ in velocity). The systematic error would certainly be expected to make the greater contribution to error-variance in representative samples of collision situations.

SUMMARY

With a modified tracking-box, *S* was presented with two targets which moved down perpendicular 'streets' toward a point of intersection. Both targets disappeared into a 'cloud' before reaching the intersection. *S*'s task on 80 trials was to indicate where the variable target would be at the instant the standard target reached the intersection. It was found that predictions made by six of the *Ss* were based upon final relative positions of the comparison-targets just before disappearance, rather than on their velocities or accelerations. No consistency was found for the predictions by the other four *Ss*.

EMMERT'S LAW AND SIZE-CONSTANCY

By ALBERT H. HASTORF, Dartmouth College, and JOHN L. KENNEDY,
The Rand Corporation

This paper reports data bearing on the points at issue between Boring, Edwards, and Young with respect to the relation between Emmert's law and size-constancy.¹ In 1940, Boring proposed a formal relationship between size-constancy and Emmert's law. He interpreted Emmert's law as referring to the increase in *apparent size* of an after-image proportional to the increase in *distance* of the projection-screen. The operations recommended by Boring for measuring *apparent size* of the after-image have been challenged by Young. Boring's method of measurement involves the use of comparison-objects,² while Young espouses direct measurement by bracketing the after-image with small spotlights.³

We propose that the controversy hangs on the characteristics of the measuring operations, and we offer the hypothesis that these operations markedly affect the phenomenon being measured. Edwards, using the comparison-method, investigated the effect of reduction-conditions on the apparent size of after-images and found that Emmert's Law did not hold.⁴ This result was not surprising to Edwards, nor should it be surprising to Young, because if Young were to repeat Edwards' experiment with the spotlight method of measurement, he would indubitably find that Emmert's Law does hold. We have confronted these two techniques of measurement with each other under reduction-conditions and find that the results come out that way.

Procedure. *O* was seated at one end of a 10 × 30-ft. room. In front of him was a 4 × 4 ft. projection-screen. To his right, at a 90° angle, was a comparison-screen (2 × 21½ ft.) which was kept at a constant distance of 3 ft. Two visual conditions were employed: (1) full binocular regard, with the room and the screens completely illuminated; and (2) monocular regard of the large screen, room illumination limited to the screens only, and view of the large screen through a reduction-screen. Judgments were made with the large screen at 20 ft.

There were four parts to the experiment. In Part 1, which dealt with size-constancy, *O*'s task was to equate the size of black stimulus-cards at the various distances and under the two visual conditions. In Part 2, in which after-images were compared, the conditions were the same, except that *O* projected an after-image on the front screen to compare with the black card on his right. To induce the after-image, *O*

* Received for publication November 4, 1955. The authors thank Jon Blum and Herbert Kleber who gathered the data described in this paper.

¹ E. G. Boring, Size constancy and Emmert's law, this JOURNAL, 53, 1940, 293-295; Ward Edwards, Emmert's law and Euclid's optics, *ibid.*, 63, 1950, 607-612; Edwards and Boring, What is Emmert's law?, *ibid.*, 64, 1951, 416-422; F. A. Young, Boring's interpretation of Emmert's law, *ibid.*, 63, 1950, 277-280; Concerning Emmert's law, *ibid.*, 64, 1951, 124-128.

² Boring, *op. cit.*, 53, 1940, 293-295.

³ Young, *op. cit.*, 63, 1950, 277-280; 64, 1951, 124-128.

⁴ Edwards, Apparent size of after-images under conditions of reduction, this JOURNAL, 66, 1953, 449-455.

fixated a 1-in., white cardboard square on a black background at a distance of 3 ft. A 25-w. flashbulb in a reflector was held somewhat below the stimulus-square and set off while *O* was fixating. Following the flash, *O* closed his eyes and waited until he had a negative after-image, at which time he opened his eyes and projected the after-image on the front screen. In Part 3, which involved Young's technique of measuring after-images with a spotlight, *O* projected the after-image, induced as described above, on the front projection-screen and bracketed it with two small spotlights, after which *E* measured the distance between the spots of light with a yardstick. In Part 4, *O* entered the experimental room when it was completely dark. His task was merely to estimate the distance of the front screen plus black card in feet to make an absolute judgment of distance. A circle of light around the card was the only light in the room. *O* gave absolute estimates, first in the dark and then under binocular regard and full lighting conditions.

Twelve undergraduate students were used as *Os* in Parts 1, 2, and 3. A new group of six *Os* was used in Part 4. One half of the *Os* went through Part 1 first, and then Parts 2 and 3, while the other group reversed this procedure. In addition, one half of each group started under the monocular conditions.

The instructions to *O* were patterned after Gilinsky's objective instructions,⁵ the aim being to establish a set favoring constancy. The instructions were:

This is an experiment to test how people judge size. Place your head in the chin rest. In front of you is a screen with a black card on it. To your right is another screen with a black card on it also. Your task is to equate these two cards so that if you placed the card on your right beside the one in front of you, or if you measured them with a ruler, they would be the same size. Look at the card in front; then turn to your right and tell me whether the card there is smaller, larger or equal to the one in front. You may look back and forth as often as you wish. Do you have any questions?

When the after-image was used, the instructions were suitably modified but the intent or set was not changed. *O* was given two practice trials with the after-image technique before data were taken. For Part 3, *O* was permitted to practice with the spotlights and was instructed to bring the spots together until the inner edges just touched the outer edges of the after-image. For Part 4, *O* was told to estimate the distance of the front screen to the nearest foot.

Results. Table I summarizes the data. The results for Parts 1 and 2 confirm Edwards' finding that the reduction-conditions affect both size-constancy judgments and after-image judgments at 20 ft.⁶ The reduction in size at 20 ft. is significant at the 1% level for both size-constancy and judgments of after-images under monocular conditions. A second important variable, which made for some variation in *N*, was the initial condition under which *O* began to make judgments. With monocular regard as the initial condition, three *Os* in Part 2 were unable to give judgments, since the smallest comparison-card was still larger than the perceived size. An acquaintance with the size of the experimental room appears to be quite important in determining *O*'s judgments of size. We also suspect that the presence of a 'real' stimulus anchors the situation, since only one of the *Os* was unable to

⁵ A. S. Gilinsky, Perception of size of objects at various distances, *Research Bulletin of Air Force Personnel and Training Research Center*, 1954.

⁶ Edwards, *op. cit.*, this JOURNAL, 66, 1953. 449-455.

make judgments under monocular regard in Part 1. Part 3, utilizing Young's method of measurement, showed no difference between reduced and full binocular conditions. Part 4, conducted with six new *Os*, demonstrated clearly that the combination of unfamiliar room, complete darkness except for a circle of light around the stimulus, and monocular regard through reduction-screen does result in significant reduction of apparent distance when absolute judgments are required.

Conclusions. These data support Boring's proposed relationship between size constancy and Emmert's Law, and confirm Edwards' results on the size of after-images under reduced conditions of observation. Our data also support Young, in

TABLE I
JUDGMENTS OF SIZE AND DISTANCE*

Part	Mean size (in.)		Mean distance (ft.)	
	binocular	monocular	binocular	monocular
1	6.6	5.3		
2	5.7	4.7		
3	6.8	6.8		
4			20.5	13.4

* The size of the standard in Part 1 was 6.66 in., as was the predicted size in Parts 2 and 3. The distance was 20 ft.

that his spotlight technique for measuring apparent size of the after-image seems to be independent of reduction-conditions, that is, yields Emmert's law under reduction-conditions. It can be argued that these results are precisely what one would expect if one makes the assumption that Boring's technique for measuring apparent size is 'psychological' whereas Young's is 'physical.' There also is some evidence that the act of measuring with the spotlights in a darkroom provides a cue (probably movement-parallax) to the distance of the projection-screen. For example, *Os* increase their estimates of the apparent distance of the screen under reduced conditions after they have been permitted to explore the screen with the spotlights as they would in measuring an after-image. If Young's technique has a built-in distance-cue, it is no surprise that his measurements are unaffected by 'reduced' conditions, whereas Boring's are. The moral of this interpretation is that psychologists, as well as physicists, need to concern themselves about methods of measurement that interact with the phenomenon being measured.

PROACTIVE AND RETROACTIVE ASSIMILATION IN THE SUCCESSIVE COMPARISON OF LOUDNESS

By HENRY GLEITMAN, Swarthmore College

The processes underlying successive comparison have often been thought continuous with those involved in memorial functions. In one sense, of course, the continuity is clear enough. If a present stimulus is to be compared with an earlier one, the retained effects of the earlier stimulus must somehow come into play, but the parallel has been extended beyond this. It has been suggested that the successive comparison is a direct means of dealing with memory-functions, capable of distinguishing between different theories of memory and forgetting, and perhaps even preferable to the more orthodox techniques by virtue of its greater simplicity.¹ In particular, emphasis has been placed on the close formal parallel between the experimental designs of retroactive inhibition and of successive comparison with interpolated stimulation. In the memory-experiment, the effect of interpolated activity is determined by comparison with a control in which the intervening interval is empty, and a decrement in retention usually is involved. In the psychophysical situation, the effect of interpolated stimuli upon judgment is determined in a similar manner, typically leading to a constant error in the direction of the interpolated stimulus which has been described as 'assimilation.'²

The present study is an attempt to test some further implications of the supposed parallelism between the two areas under discussion. Assuming this parallelism, one would expect that interpolative effects and memorial interference would show similar functional relationships to other variables. The variable here chosen has been the order in which materials are presented, *i.e.* the differential effect of proactive and retroactive inhibition. Such a differential effect has been demonstrated repeatedly in the area of memorial interference: When the time intervals between original learning and test are relatively short, proactive inhibition is weaker than retroactive inhibition.³ The purpose of this paper is to discover whether a similar relation holds in the area of successive comparison: Is the 'assimilative' effect due

* Received for publication November 20, 1955. This study was performed in the laboratories of the Department of Psychology of Cornell University. The writer wishes to thank the members of that department for their generous coöperation.

¹ C. C. Pratt, The law of disuse, *Psychol. Rev.*, 43, 1936, 83-93.

² J. P. Guilford and D. G. Park, The effect of interpolated weights upon comparative judgments, this JOURNAL, 43, 1931, 589-599; Otto Lauenstein, Ansatz zu einer physiologischen Theorie des Vergleichs und der Zeitfehler, *Psychol. Forsch.*, 17, 1932, 130-177; J. G. Needham, Interpolation effects with different time-intervals, *J. exp. Psychol.*, 18, 1935, 767-773; Pratt, The time-error in psychophysical judgments, this JOURNAL, 45, 1933, 292-297.

³ A. W. Melton and W. J. Von Lackum, Retroactive and proactive inhibition in retention: evidence for a two-factor theory of retroactive inhibition, this JOURNAL, 54, 1941, 157-173; J. A. McGeoch and B. J. Underwood, Tests of the two-factor theory of retroactive inhibition, *J. exp. Psychol.*, 32, 1943, 1-16; Underwood, Retroactive and proactive inhibition after five and forty-eight hours, *ibid.*, 38, 1948, 29-38.

to a stimulus interpolated between standard and variable greater than that caused by a stimulus which precedes the standard?⁴

Method. The present experiment is patterned after the paradigms for the study of retroactive and proactive inhibition, but adapted to the situation involved in successive comparison. Three conditions were employed: control, retroactive, and proactive interference. In the control condition, a variable stimulus was compared with a previously presented standard. In the retroactive condition, a more intense stimulus was interpolated between standard and variable. In the proactive condition, the same intense stimulus *preceded* the standard.

The stimuli used were pure tones, at a fixed frequency of 1000—. Each tone was presented for 0.5 sec. Stimulus *B* (preceding for the proactive, interpolated for the retroactive condition) was set at 80 db., and the standard at 60 db. The interval between Stimulus *B* and the standard was always 2 sec., regardless of whether *B* preceded or followed the standard. To lessen the possible effects of assimilation across comparison-pairs, a 20-sec. interval was introduced after every judgment.

Since differences between the proactive and retroactive conditions might be a function of the interval between standard and variable—as they are functions of the interval between original learning and recall in the memory experiment—two such intervals were employed, 4 and 6 sec. Thus there were in all conditions, the three main conditions each being further split to allow for two standard variable intervals.

Apparatus. The stimuli were generated by a Hewlett-Packard audio-oscillator (Model 200 CD) and fed into calibrated PRD-8 binaural earphones. The click was eliminated by a system of filters. All stimuli were generated by the same oscillator, across the output of which three resistances were placed. The first two of these were fixed for Stimulus *B* and for the standard stimulus, while the third was variable and regulated the intensity of the comparison tone. The duration of the stimuli and the intervals between them were regulated by a series of timing relays connected to appropriate shunts.

Procedure. Six *O*s took part in the experiment, four men and two women, all graduate students. They had never served as *O*s in an auditory experiment and did not know the purpose of the present experiment. Each *O* was subjected to all six conditions. To factor out any possible long-range effects of practice, the conditions were presented in different order for each *O* with every condition appeared once in every one of six possible positions. Each condition was presented in a separate session, the sessions being scheduled for the most part on alternate days.

The method of constant stimuli was used. There were 11 comparison-stimuli, ranging from 54 to 67 db. The instructions were to judge the comparison-stimulus as louder or softer than the standard; no equal judgments were allowed. Each comparison-stimulus was presented 10 times in random order. The average duration of an experimental session was 1 hr.

⁴ Assimilative effects due to preceding stimulation have been demonstrated in previous studies (S. W. Fernberger, Interdependence of judgments within the series for the method of constant stimuli, *J. exp. Psychol.*, 3, 1920, 126-150; Louis Long, A study of the effect of preceding stimuli upon the judgment of auditory intensities, *Arch. Psychol.*, 30, 1937, (No. 209), 1-57).

Results. Individual psychometric functions were fitted for each *O* under each condition, yielding 36 functions in all. The data were fitted to normal ogives, using the Müller-Urban weights. On the basis of these curves two constants were calculated for each of the functions—the point of subjective equality, (*PSE*), and the *SD*. Table I shows the mean *PSE* and *SD* for each condition.

An inspection of the *PSE* values indicates both a proactive and a retroactive assimilation effect—the introduction of the 80-db. stimulus leads to a positive time-error in both experimental conditions. The retroactive effect is by no means stronger than the proactive; if anything, the opposite holds true. An analysis of variance yielded no significant effect of time-interval between standard and variable, nor was there a significant interaction between time-interval and the main conditions. The main effect was significant ($F = 6.98$, $P < 0.01$), but the difference between

TABLE I
MEAN *PSE* AND *SD* FOR EACH CONDITION

	Proactive		Retroactive		Control	
	4 sec.	6 sec.	4 sec.	6 sec.	4 sec.	6 sec.
<i>PSE</i>	62.26	61.76	60.84	62.48	59.68	59.77
<i>SD</i>	6.03	7.11	4.91	5.07	3.65	4.26

proactive and retroactive stimulation was not ($t = 0.53$)—proactive and retroactive assimilation effects appeared to be of equal strength.

The effect of interpolated stimulation upon successive comparison has usually been studied in connection with change in the *PSE*—the extent to which extraneous stimuli lead to over- or under-estimation of the standard. It is possible also that change in precision of judgment may be found. Table I shows that extraneous stimulation produced a marked increase in the *SD* of judgment. An analysis of variance yielded a significant main effect ($F = 9.76$, $P < 0.01$), and the difference between the proactive and the retroactive situation also proved to be significant ($t = 2.62$, $P < 0.05$). The difference between the control and the retroactive condition was not significant ($t = 1.73$, $P < 0.20$). The effects of time-interval and of the interaction between time-interval and main conditions were not significant.

Summary. A third and more intense stimulus was found to affect the successive comparison of two less intense stimuli, both when it preceded and when it followed the standard. Adding the third stimulus led to an over-estimation of the standard and a reduced precision of judgment. Neither of these effects was greater for the retroactive than for the proactive condition. The results suggest that the analogy between the interpolation experiment in successive-comparison and retroactive-inhibition studies is relatively superficial. The difference between proactive and retroactive inhibition is well established and rather general. Since the same is not true of the psychophysical situation, the proposed parallelism between the two areas is suspect.

FREQUENCY-CHARACTERISTICS OF SUCCESSIVE WORD-ASSOCIATION

By EPHRAIM ROSEN and WALLACE A. RUSSELL, University of Minnesota

It has commonly been assumed that the cultural frequency of a response in free association is an index of the strength of that response among the individuals in the population sampled. For example, since the stimulus-word *chair* is responded to with the word *table* by about 50% of a normative group as contrasted with 20% for the response of *sit*,¹ it is often assumed that in any given individual's repertoire of responses to *chair*, *table* has a probability two and one half times that of *sit*. Evidence for the validity of this inference has been reported by Schlosberg and Heineman, who found a correlation of -0.80 between cultural frequency and log reaction-time²

More evidence is provided by the present study, which was designed to investigate the relation between the hierarchy of responses obtained by eliciting a single response across many individuals (cultural frequency) and the hierarchy of responses obtained by getting successive associations within individuals.

Method. The standard conditions of free association were so modified that each S was required to give two successive responses to a list of stimulus-words. The instructions were as follows:

I'm going to read a list of 100 different words to you. I will say each word twice (in succession), and each time I say it I want you to respond with the first word you think of. The second time I say a word, try to respond with a different word from the one you used the first time I said the word. For example, if I say *a* you might say *b*; then when I say *a* again, you might respond with *c* or *x*, but don't say *b* again unless nothing else comes to mind. Do you understand?

The stimulus-words used were taken from the Kent-Rosanoff list.³ The present report covers responses to the first 50 words of this list. The words were presented orally by E to each of 63 Ss, (men, undergraduate college students) who responded orally. The entire test with each S was recorded and the responses were transcribed.

Results. Three types of analysis were employed to determine the frequency-characteristics of the successive associations obtained in the experimental situation. All the analyses were oriented toward testing the hypothesis that the individual tends to respond in successive association with the same sequence that composes the cultural frequency-hierarchy for single associations.

* Accepted for publication December 8, 1955. This study is part of a larger series of studies of verbal behavior sponsored by the Office of Naval Research (Contract No. N8 onr-66216) under its policy of encouraging basic research. The authors were assisted in the analysis of results by Neal Johnson.

¹ W. A. Russell and J. J. Jenkins, *The Complete Minnesota Norms for Responses to 100 words from the Kent-Rosanoff Word Association Test*, Technical Report No. 11, ONR, Contract No. N8 onr-66216, University of Minnesota, 1954.

² Harold Schlosberg and Charles Heineman, The relationship between two measures of response strength, *J. exp. Psychol.*, 40, 1950, 235-247.

³ G. H. Kent and A. J. Rosanoff, A study of association in insanity, *Amer. J. Insanity*, 67, 1910, 37-96, 317-390.

(1) If this hypothesis is correct, then the second response in successive association should have a cultural frequency lower than that of the first response. For all of our stimulus-words, the cultural frequencies of the first and second responses were compared. The range of the average cultural frequency of response to each stimulus-word was 2.00-77.34%, with a mean of 22.72%, for the first response; and 1.04-17.42%, with a mean of 7.32%, for the second response. The difference between the means (15.40%) is highly reliable, since there were no reversals in direction; for every one of the 50 words the average cultural frequency of the second response was lower than that of the first.

(2) It might be objected that the large mean-difference was due solely to the fact that Ss who respond first with the cultural primary, that is with the most frequent response in the cultural norms, must inevitably give a less frequent second response. Another test of the major hypothesis, not as open to this criticism, would be to determine whether S tends to produce a second response with a lower cultural frequency than that of his *own* first response. This test is more stringent than the first one, for among Ss who fail to give the culturally most frequent response first there is ample opportunity for the cultural frequency of the second response to be either higher or lower than that of the first. When this test was made, however, it was found that 41 (65.0%) of the Ss gave a second response with a lower cultural frequency than the first; 18 (28.6%), gave a second response with a higher cultural frequency than the first; and 4 (6.4%), gave two responses of the same cultural frequency. Of the latter 4 Ss, 2 violated the instructions by giving the same response twice. Even when the Ss who gave a cultural primary as the first response are subtracted from those conforming to the hypothesis, the first category still contains a larger number of Ss (19) than either of the other two groups.

A typical example of the distribution of Ss within these patterns, for a single stimulus-word, is the following: To the stimulus-word *mutton*, for which the primary response is *lamb*, 41 Ss gave second responses with lower cultural frequencies than their first responses. Of these, 21 gave a primary followed by a non-primary, and 20 a non-primary followed by a less frequent non-primary. For 18 Ss, the cultural frequency of the second response was higher than that of the first, while for 3 others the two responses were equal in frequency.

(3) Finally, a very stringent test of the hypothesis would be to determine if the most frequent second response in successive association was actually the second most frequent response in the cultural norms. Furthermore, to the extent to which the second responses conform to the sequence found in the cultural norms, a positive rank-order correlation would be expected between the hierarchy of second responses and the cultural frequency-hierarchy.

To test these predictions, the 15 stimulus-words with the culturally most frequent secondary responses were selected. The most frequent second response in the experimental group coincided with the normative cultural secondary in 8 of the 15 cases. In the remaining cases, the second response was the same as the culturally *third*-ranking response twice, and the same as the cultural primary five times. Thus, in the majority of cases the prediction was borne out. Furthermore, in four of the five cases in which the cultural primary was the most popular second response, the cultural primary was given most frequently both as the first

and second response. It would appear then that for these four words the cultural primary is of such strength that it would be necessary to have more than two successive associations to proceed far enough down the response-hierarchy to reach the cultural secondary.

For the same 15 stimulus-words, rank-order correlations were computed between the rankings of all responses occurring more than once in the second responses of the 63 Ss and the rankings of these same responses (relative to each other) in the cultural norms. The number of such response-words varied from 3 to 12 in the 15 correlations, all of which were positive, ranging from 0.38-0.98 with a median of 0.75. While little statistical confidence can be placed in any one of these correlations due to the small number of cases involved, the positive direction of all 15 is striking.

Conclusion. The technique of successive association provides support for the assumption that the cultural frequency of an association may be taken as an index of the strength of the response.

THE EFFECT OF EFFORT AND TRAINING-TEST SIMILARITY ON RESISTANCE TO EXTINCTION

By REED LAWSON and AARON J. BROWNSTEIN, University of Missouri

In Hullian theory, work (reactive inhibition) is a primary determinant of resistance to extinction,¹ but the data of a recent experiment contradict that assumption.² Furthermore, afferent theories have been proposed which entirely ignore the role of work *per se*—change in the effortfulness of response, whether upward or downward, may facilitate extinction, but only to the degree that discrimination between training and extinction series is facilitated³ or the number of conditioned stimulus-elements is reduced.⁴ For Hullian theory, too, any change in stimulating conditions tends, by the principle of stimulus-generalization, to reduce resistance to extinction, but, in the case of reduction in effortfulness, the effects of stimulus-generalization and reactive inhibition are opposed, and the outcome cannot be predicted. For the alternative theories, reduction in effortfulness can tend only to facilitate extinction. An empirical test is provided by the experiment to be reported. Rats were trained to traverse a runway containing several hurdles, which, for half the Ss, were removed during extinction. Removal of the hurdles was presumed to reduce the effortfulness of response (defined in molar fashion) and simultaneously to increase the discriminability of training and extinction situations.

Method. The Ss were 26 albino rats, 14 males aged 119 days, and 12 females aged 153–162 days. Three males were discarded for failure to run, and one for illness.

The apparatus employed was a straight runway, 56 in. long, 4 in. wide, and 4¾ in. deep, with a starting box 8 in. in length. One guillotine door separated the starting box and runway, another was 8 in. from the end of the runway. The end-box was black, 1 ft. long, and of the same width and depth as the runway. There were three hurdles, spaced 1 ft. apart and 1 ft. from each of the guillotine doors. The hurdles, 2 in. high, with an upper surface 2 in. across and 4 in. wide, were painted black. A 25-w. lamp 22 in. above the floor of the runway-segment beyond the second guillotine-door provided the only illumination in the room. The runway faced into a small alcove, the three walls of which were covered by black cloth.

On each rewarded trial, S received two 40-45 mg. pellets of food in a black glass cup; the daily supplementary ration was dry mash. Water was always available in the living-cages, which were situated in the experimental room. An air-conditioner

* Received for publication December 26, 1955.

¹ C. L. Hull, *Principles of Behavior*, 1943, 258-303.

² J. L. Maatsch, H. M. Adelman, and M. R. Denny, Effort and resistance to extinction of the bar-pressing response, *J. compar. physiol. Psychol.*, 47, 1954, 47-50.

³ D. W. Tyler, E. C. Wortz, and M. E. Bitterman, The effect of random and alternating partial reinforcement on resistance to extinction in the rat, this JOURNAL, 66, 1953, 57-65; M. E. Bitterman, W. E. Feddersen, and D. W. Tyler, Secondary reinforcement and the discrimination hypothesis, this JOURNAL, 66, 1953, 456-464.

⁴ W. K. Estes, Statistical theory of distributional phenomena in learning, *Psychol. Rev.*, 62, 1955, 369-377.

maintained a room-temperature between 72-78°F. and helped also to mask extraneous noise.

The Ss were placed on a 22-hr. deprivation-schedule and, for 5 days, adapted to handling and given social adaptation in the runway, with and without the hurdles. On each of the next 5 days, each S was given one rewarded trial. On each of the next 2 days, 6 massed trials were given; Trials 2, 4, and 6 were not rewarded. On the final training day, there were 6 additional trials, of which Trials 1, 3, and 5 were not rewarded. On nonrewarded training trials, the food-cup was present but empty.

On the extinction trials, the 11 Ss of Group I had the food and food-cup removed; the only additional change for the 11 Ss of Group II was the absence of the hurdles. All Ss were given massed extinction trials to a criterion of 2 successive failures to enter the end-box within 2 min. after the starting door was raised. From 15-20 hr. after the extinction-series, Ss were fed for 2 hr., and 20 hrs. later were given two spontaneous-recovery trials on which the extinction-conditions were repeated.

Results. Group I (tested with hurdles) was superior in resistance to extinction. The distribution of trials to extinction for the two groups was as follows: for Group I: 10, 42, 53, 57, 72, 74, 88, 115, 126, 168, 188; for Group II: 7, 9, 20, 26, 27, 33, 35, 44, 101, 107, 146. By the Mann-Whitney rank-test,⁵ the null hypothesis may be rejected at the 5% level. The possibility that these results might be due to uncontrolled variables was not supported by several control analyses. The average $\log_{10}$ running-time curves of the two groups did not differ in level or in slope on the final training day, nor did the two groups differ significantly in age, in weight at start of training, in weight at extinction, or in percentage of weight lost.

Different recovery-phenomena were shown by the two groups in the final tests. The average $\log_{10}$ running time of Group II was lower on the first recovery-trial than on the first extinction-trial, while the opposite was true for Group I (a result more typical of spontaneous recovery). The difference between the two changes was statistically significant ($F = 5.10$, $p < 0.05$).

The mean $\log_{10}$ running-time curves of the two groups on the first six extinction-trials differed neither in level nor in slope. The level-comparison is spurious, of course, because of the nature of the apparatus; an uninterrupted run in the presence of hurdles would perforce be slower than a run in the absence of hurdles. The lack of a difference in slope is of some interest, however, and will be discussed below. In terms of actual performance on the spontaneous-recovery trials, Group II was faster than Group I on the first trial, slower on the second; this trials $\times$ group-interaction was statistically significant ($F = 8.068$, $p < 0.025$ for 1 and 20 *df.*).

Discussion. While these data do not rule out work-produced inhibition as a factor in extinction, they suggest that to regard it as a primary determinant of resistance to extinction is, at least, cumbersome. For example, it is possible that the generalized reaction-potential of Group II was so low as to require little inhibition to bring it below the excitatory threshold. By the same token, however, the conditioned

⁵ H. B. Mann and D. R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, *Ann. Math. Stat.*, 18, 1947, 50-60.

inhibition presumably developed during the massed training trials would also have been present in a much smaller amount in Group II than in Group I. There also are other complications. The major point to be made is that if a Hullian account of these results must ultimately place primary emphasis on the effects of stimulus-generalization, such an account becomes similar, except in complexity, to those of the 'discriminative' and 'statistical' theories.

These data are interpreted equally well by the two stimulus-oriented theories, but the theories might yield different predictions if the amount of training in this situation were greatly increased. It appears that the discrimination hypothesis still would predict the present outcome. In Estes' view, on the other hand, continued training might condition enough stimuli not associated with the hurdles to the running response that, ultimately, removal of the hurdles might not appreciably reduce the proportion of available conditioned stimuli; the probability of response in Group II thus would be little reduced.

An aspect of the present results independent of the preceding considerations relates to spontaneous recovery. Group II apparently improved its performance during a series of nonrewarded trials; the Ss ran significantly faster on the first recovery-trial than on the first extinction-trial. It might be proposed that the first exposure of the empty runway produced external inhibition, thus making the first extinction-time a distorted indicator of the strength of responding, but there are several objections to this view. Classically, external inhibition is produced by the presence of new stimuli, not the absence of old ones. Furthermore, over the first six extinction-trials the running-times of the two groups increased at the same rate; there was no evidence of recovery from presumed external inhibition, which might be expected to occur. Actually, there seem to be no criteria of the presence or absence of external inhibition under such circumstances; with one exception, all the Ss of Group II did eventually respond on each of these trials. It should be noted, too, that all Ss originally were given two adaptation-sessions in the empty runway to minimize the possibility of novelty, or fear-producing, effects when it was presented during extinction.

An hypothesis which suggests itself is that the Ss of Group II were acquiring a new 'skill' during the extinction-series. Literally, they were learning to run in an alley without hurdles. This learning probably involved the elimination of incipient hurdle-jumping tendencies that were now incompatible with an efficient run. To be sure, this skill was a useless one; as the results for the second spontaneous-recovery-trial showed, there was little motivation to employ it in this situation.

Summary. The relative influence of effort and of similarity between training- and extinction-conditions upon resistance to extinction was studied. Rats were trained to traverse a runway containing three hurdles, and half the Ss were given extinction-trials with the hurdles absent. The Ss for which the hurdles remained in place showed significantly greater resistance to extinction, although they were presumably expending more effort per trial. On a subsequent spontaneous-recovery trial under the former extinction-conditions, Ss for whom the hurdles were absent actually exceeded their performance at the beginning of the extinction-series.

UNRECOGNIZED WEAKNESS AND COMPENSATORY LEARNING

By RICHARD K. OVERTON and W. LYNN BROWN, University of Texas

The present paper reports a test of the hypothesis that a permanent physiological weakness may elicit compensatory learning even though the *S* has never been aware of the weakness. The hypothesis is similar to the late Alfred Adler's concept of "organ inferiority."¹ Heibreder and others have shown that man may employ cues of which he is not conscious. Heibreder's *Ss* were, however, consciously motivated to solve a known problem. The present study, in contrast, involves *Ss* who say they are not aware that a physiological weakness has presented them a problem.²

Method: Subjects. The *Ss*, 74 in number (50 men and 24 women), were students in undergraduate psychology and classes in Russian.

Equipment. Pseudo-isochromatic plates, products of the American Optical Company,³ were used for testing red-green color vision. They were employed according to company specifications except that the standard source of light used in the experiment provided slightly less illumination than the source specified. The weaker source presumably made the test slightly more difficult for persons with borderline deficiency in color perception.

Procedure. Each *S* was told: This is just a test for color vision. Have you ever suspected that you might be a little color blind? After *S*'s response had been recorded, the test plates were administered. The *E* then said: Now I'm going to ask you a question about the traffic lights here in Austin. On most traffic lights, is the red light or the green light on top? Time required by *S* to answer was recorded. Timing was begun at the word 'top' and ended when *S* named a color. All traffic lights in Austin and most other cities have the red light above the green.

Results. Two *Ss*, both men, said they had previously suspected deficient color vision in themselves. Data from these two *Ss* were not used.

Results obtained from the other men are presented in Table I. The table indicates that *Ss* with some color weakness tended to give quick, correct answers to the question about the traffic light. In other words, there was a tendency for defective color vision to elicit compensatory learning even in *Ss* who were not aware of their slight deficiency.

The significance of this tendency may be determined by a Chi-squared analysis of Table I. The table yields, after the consolidation of the three right columns, a Chi-square of 9.626 with two degrees of freedom. This result permits acceptance of the hypothesis under investigation at the 1% level of confidence.

Table I also presents data obtained from the 24 women, which points similarly

* Received for publication October 17, 1955.

¹ A. Adler, A study of organ inferiority and its psychical compensation, *Nerv. ment. Dis. Monogr.*, 24, 1917, 1-86.

² Edna Heibreder, The attainment of concepts: III. The process, *J. Psychol.*, 24, 1947, 93-138; Edna Heibreder, M. D. Bensley, and M. Ivy, The attainment of concepts: IV. Regularities and levels, *ibid.*, 25, 1948, 299-329.

³ American Optical Company, Pseudo-isochromatic plates for testing color perception, 1946.

to a positive correlation between color weakness and learning. Consolidation of the two right columns and the two bottom rows of Table I into one row and one

TABLE I
COLOR WEAKNESS AND RESPONSE TO QUESTION ABOUT TRAFFIC LIGHT
Number of color plates missed

	Men				Women		
	0	1	2	3 or more	0	1	2 or more
Correct in <3 Sec.:	5	5	5	7	2	5	5
Correct in 3-10 Sec.:	8	6	0	0	4	1	1
Correct in >10 Sec., or Incorrect:	9	2	0	1	5 11	1 7	0 6

column yields a Chi-square of 6.226. With one degree of freedom, this Chi-square is significant at the 2% level of confidence. Both men and women give results which support the hypothesis advanced.

APPARATUS

A MAZE FOR PLANARIA

By N. H. PRONKO and LOUIS WEHRENBURG, JR., University of Wichita

After many disappointments, a planarian maze has been devised which works satisfactorily for these animals. It has the added advantage of requiring the minimal handling of Ss, in the particular problem for which it was designed. Animals need only to be put into the maze and removed from it. Handling them between trials is unnecessary.

The entire maze was machined out of aluminum. It consists of a housing (E) and an over-all plastic cover which has been found unnecessary. With only the alleys

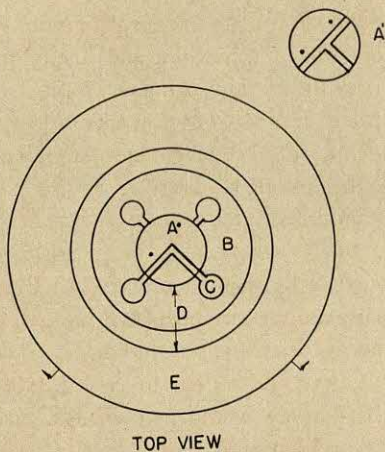


FIG. 1. PLANARIAN MAZE

A = training maze, R or L turn, stationary but removable; A' = test-maze ('T'), R or L turn, stationary but removable; B = entrance and goal compartments, may be rotated; D = plastic cover; E = housing.

and starting and goal compartments moist, the animals do not venture out of the uncovered maze alley. An animal is put into C, and a light flashed on it. Immediately, it is activated and after a left turn, it winds up at the compartment at the other end of the maze. The movable platform is then rotated to the C position and the animal is ready for another trial. This can be repeated any number of times. Right turns are handled in the opposite direction, with C as the goal compartment.

Removal of A and insertion of A' permits testing of the efficacy of training before or after transection of the animal.

The chief advantage of this maze over previous ones is in its preventing the Planaria from travelling on the vertical walls of the maze. The shallowness of the maze forces the animal to locomote on the floor exclusively.

NOTES AND DISCUSSIONS

TECHNICAL AND SCIENTIFIC COMMUNICATION: A REPLY TO CALVERT

In the September 1956 issue of this JOURNAL, E. S. Calvert, of the Royal Aircraft Establishment in Britain, offers some rather blunt advice. If psychologists wish to make a useful contribution to aviation, he says, they should make themselves familiar with his theory of the visual basis for landing an aircraft—the streamer theory—and they should “adopt in the future a less muddled terminology.”¹ The occasion for this advice was a paper in the September 1955 issue, for which I was primarily responsible.² It gave mathematical expression to the concept of *motion perspective* as a basis for the control of aerial locomotion, and distinguished this concept from the classical notion of motion parallax employed in physiological optics and space perception. It is obvious that theories of visual stimulation arouse more than academic interest when they apply to that moment of hazard in modern life—the aircraft landing.

Now the supposedly muddled term *motion perspective* has been in print for psychologists since 1947,³ whereas the term *streamer* and its variants could hardly have been read by them before 1954.⁴ Even if it had previously been in common use among engineers and pilots, psychologists could not be expected to know it. Scientific terminology is established by a body of scientific literature. Calvert fails even to give a reference to his own publication in telling us that we would be well advised to know his way of thinking about these problems. My paper should have referred to this paper of his but I missed it. Even more, he should have taken account of my published theorizing, but he did not. The only cause for complaint is a miserable failure in scientific and technical communication.

Calvert's achievements abroad have been in the field of airport runway and approach lighting. He analyzed the dynamic field of view of a pilot coming in for a low-visibility landing with perspective diagrams—hence

¹ E. S. Calvert, Technical terms in science and technology, this JOURNAL, 69, 1956, 476-479.

² J. J. Gibson, Paul Olum, and Frank Rosenblatt, Parallax and perspective during aircraft landings, this JOURNAL, 68, 1955, 372-385. Also Gibson, The optical expansion pattern in aerial locomotion, *ibid.*, 480-484.

³ J. J. Gibson (ed.), *Motion Picture Testing and Research*, AAF Aviation Psychology Research Report No. 7, 1947, Ch. 9, pp. 219 ff. Also *The Perception of the Visual World*, 1951, Ch. 7.

⁴ Calvert, Visual judgments in motion, *J. Instit. Navigation*, 7, 1954, 233-251 (published by the Royal Geographical Society, 1 Kensington Gore, London, S. W. 7).

the "streamers." He also engineered practical reforms in airport lighting which have undoubtedly saved lives. His theory of optical information during landing is similar to my own except for terminology and formulation; but his first reports of the perspective analysis, after the war, were classified for reasons of military security—what reasons I cannot imagine. They could be read, after delay and difficulty, but not kept and not referred to in print. This is the sure and certain way of strangling scientific communication. Matters are just as bad or worse over here. Interesting research in this country on optical information for flight, even today, is being carried out without permission to publish. It will become ingrown and provincial as a result.

Calvert's concept of streamers in the field of view and my concept of motion perspective are based on the same facts of optical geometry, whether the latter term is a fortunate one or not. The aeronautical profession has been slow to accept these facts, so we ought to concentrate our effort there and not waste it in terminological disputes. I put it this way: the pattern of the light projected from an environment to a point in space is governed by the geometry of perspective. Hence it is useful to say that the *flow*-pattern of the light projected from an environment to a *moving* point in space is governed by a geometry of *motion* perspective. The words make no difference so long as this special geometry can be agreed upon.

I am convinced that there is no inconsistency between the mathematics employed by Calvert to describe streamer-patterns and that employed by me and my collaborators to describe flow-patterns. During my visit to England last year he and I spent most of a day discussing the optical information for the control of an airplane which enters a pilot's eye. If this is expressed in terms of angular coördinates, one obtains a general mathematical statement of the "law of motion perspective,"⁵ which is, of course, neat; Calvert wishes to express the information in terms of a picture-plane. It is only a matter of preference. The questions which we disagreed on and argued about were psychological, not mathematical. Calvert has convictions about the *specific knowledge* a pilot has to have in order to make a good landing and the *particular judgments* he needs to make during the process. For my part, having studied the introspections of pilots during the war and tried it myself, I would just as leave ask a bee how it lands on a flower as a pilot how he lands on a runway. Neither one knows. Nevertheless both do. The effective variables in the flux of stimulation are not necessarily the set of 'cues' which a pilot is aware of. Calvert's list of cues

⁵ Gibson, Olum, and Rosenblatt, *op. cit.*, 379.

and judgments is better than the venerable textbook list of cues for depth perception, but it is still akin. The psychological problem of the control of locomotor action by optical stimulation is not solved by such a list.

Calvert is an engineer who wishes to prevent accidents. I am a psychologist who wishes to understand visually guided locomotion. When he argues that the proper meanings of terms like *parallax*, *perspective*, *orientation*, and *texture* are to be found in any technical dictionary and that any effort to change these meanings only muddles thinking and confuses the reader, I have to answer that new concepts demand changes in the meanings of old terms. A good theory justifies a great deal of linguistic inconvenience; and in the long run a good theory will prevent more accidents than a muddled theory.

JAMES J. GIBSON

Cornell University

THE INFLUENCE OF BACKGROUND ON THE PERCEPTION OF CONVEXITY AND CONCAVITY

Certain types of figures, which are characterized by a gradient of light and shade, yield a perception of convexity when held in one position and a perception of concavity when rotated 180° . In this regard the 'monitor illusion' is probably best known.¹ When this picture is held in an upright position the large elliptical figures are typically seen as concave and the rivets are seen as convex. Reverse depth-effects are obtained when the picture is rotated 180° . The distribution of light and shade across the picture in its normal orientation is such as would be produced by an overhead light source. One explanation of this effect has been in terms of learning.² This explanation as usually stated involves the following assumptions: (1) the direction of light in the individual's experience has been from above; and (2) the differential cues of light and shade produced by convex and concave surfaces are appropriately interpreted and perceived. As an alternative to the first of these assumptions, the present report considers the possible influence of the background of the figure. The writer noted that the four shaded circles on a white background, which are to be found in Gibson's text on perception, did not lead to a reversal of depth with rotation.³ This result suggests that in other figures, such as the

¹ S. W. Fernberger, *Elementary General Psychology*, 1936, 88.

² J. F. Dashiell, *Fundamentals of General Psychology*, 1937, 450; E. H. Hess, Development of the chick's responses to light and shade cues of depth, *J. comp. physiol. Psychol.*, 43, 1950, 112-122.

³ J. J. Gibson, *Perception of the Visual World*, 1950, 95.

'monitor illusion,' the number of reversals in depth would be significantly decreased by surrounding an elliptical figure by a white background.

The large elliptical figure in the upper left hand corner of the picture of the monitor "Lehigh" was given a white surround which extended to the horizontal edge above and to the vertical edge to the left. The picture was rephotographed to eliminate a visible contour. This picture with the critical figure on it and the original picture (without any white surround) were presented to 50 Ss who perceived the expected depth of the elliptical figure on the original (control) picture. Both the order of presentation and orientation of the two pictures were systematically changed. Thirty-two of the 50 Ss reported seeing the *critical* figure as convex in both orientations. We can conclude that the white surround significantly influences the perception of the critical figure in the direction of convexity.

Assumption 1 above is, apart from the results of this study, of uncertain validity since it may be contended that light, especially in the typical home, is multi-directional rather than uni-directional. Hess' experiment, which was designed to demonstrate the causal role of variation in the direction of light in the formative period of groups of chicks, is relevant in this connection.⁴ This experiment, however, surely an important step forward in dealing with problems of this kind, is not decisive because during the training period the chicks might have learned to associate the direction of shadow cast by the grain with pecking. This learned association would then be transferred to the photographs of grain.

Queens College

NICHOLAS PASTORE

AUTONOMOUS CHANGES IN THE MEMORY-TRACE

Walker and Veroff have recently provided evidence to support the hypothesis that memory-traces undergo a form of autonomous stabilization with the passage of time as measured by the reproduction of perceived form.¹ The study demonstrates that the effects of dynamic activity within the individual memory-trace can be measured by the reproductive method when the error introduced by the act of reproduction itself is taken into account.

Several years ago, the writer was convinced that research failure to detect autonomous, structural changes in the memory-trace was due to interference of waking activities on the trace system. An unpublished pilot study was

⁴ Hess, *op. cit.*, 122.

¹ E. L. Walker and J. Veroff, Changes in the memory-trace for perceived forms with successive reproductions, this JOURNAL, 69, 1956, 395-402.

performed to test this hypothesis. The rationale for the experimental design was derived from the Jenkins and Dallenbach study of the retention of nonsense syllables after an interval of sleep and waking.² Unfortunately, the study lacked adequate control of the sleep interval; nevertheless, the results indicated that autonomous changes in the memory-trace can be readily detected when the interfering activity in the retention interval is held at a minimum.

The purpose of this note is to suggest that future research in this area can be accomplished by systematic measurement of autonomous changes following an interval of sleep. Adequate control of subjects' sleeping behavior may present a serious problem but this obstacle is not insurmountable with our present devices, for example, the electroencephalogram. Other methods for controlling interfering activity can also be suggested such as hypnosis or drugs. In accord with Walker and Veroff, "much more work needs to be done on the problem of conditions that will predispose toward each of the antithetical processes."³

Tyndall Air Force Base, Florida

JACK E. CONKLIN

AN AUDITORY AUTOKINETIC EFFECT

While considerable attention has been given to the apparent motion of a stationary point of light during prolonged inspection in the dark, little is known of analogous phenomena in other sense-modalities. Preliminary observations of an auditory autokinetic effect are here reported. Since spontaneous variation might have occurred along any auditory dimension, such as pitch, loudness, or spatial position, we began with an exploratory study which would not prejudice the issue in favor of any particular dimension.

The stimuli were relatively pure tones of 56, 560, and 5600~, generated by an audio-oscillator. The *O* was seated in a dark room, 12 ft. from a loudspeaker. For the first observations, the tones were presented once each in random order. The duration of stimulation was 5 min., and there was a 1-min. rest between stimuli. Instructions were simply to "report what you notice about the sounds." All *O*s (10 undergraduates) reported spatial displacement of the sounds, and 3 *O*s reported less striking variations in loudness and pitch as well. The movement of the 5600 ~ tone

² J. G. Jenkins and K. M. Dallenbach, Obliviscence during sleep and waking, this JOURNAL, 35, 1924, 605-612. For the effect of activity on memory-traces see also Hiroshi Mimami and K. M. Dallenbach, The effect of activity upon learning and retention in the cockroach, *ibid.*, 59, 1946, 10-36.

³ Walker and Veroff, *op. cit.*, 402.

seemed to be of greater magnitude than the movement of the lower tones.

In a second series, the method was essentially the same as before, except that *Os* were specifically instructed to listen for changes in pitch and loudness, while nothing was said about changes in position. The *Os* (5 undergraduates) all reported changes in frequency and pitch for all three stimuli, and all gave some report of spatial displacement as well.

In a third series, the three frequencies were employed once more for 20 *Os*. For half the group, a small point-source of light was placed on top of the loudspeaker. All *Os* were told that the sound would appear to move and were instructed to give a running commentary on the direction and extent of movement. The group with light and sound was instructed to report only the movements of the sound, but to watch the light carefully and to be prepared to describe its movement after the experiment with special reference to whether the light and sound had seemed to move together. The order of presentation of stimuli was counterbalanced.

Every *O* reported some movement of the sound, although there were large individual differences in amount of reported movement. The movement was in three-dimensional space. It had the same rather ethereal, floating quality which is often reported in the visual case. Although *Os* found it difficult to estimate the extent of the movement, it was relatively easy to report on the occurrence and direction of movement. In the quantification of results, a separate movement was scored for every onset of movement after a stationary period, and for every change in direction.

The number of movements reported in the 15 min. of observation for each *O* ranged from 10 to 151, with a median of 56 reported movements for both groups combined. Every *O* in the 'Light Group' showed the visual auto-kinetic effect. The number of movements was somewhat greater in that group, but not significantly so. Without the light, there was a tendency again for the number of movements to increase with frequency, while no such tendency appeared in the presence of light. The remarks of the *Os* suggest that the two kinds of movement frequently occurred along the same *line*: for some *Os* the light and sound tended to move in the same direction, while for others they tended to move in opposite directions.

In these experiments *O*'s head was not fixed. Further observations with four *Os* show, however, that a marked auditory autokinetic effect may appear when *O*'s head is fixed.¹

ALFRED C. BERNARDIN

University of Colorado

HOWARD E. GRUBER

¹ We are indebted to Dr. Harry Hurwitz of Birkbeck College, University of London, for his help in making these observations.

EDITORIAL ANNOUNCEMENTS

We are pleased to announce that the following have accepted membership upon the JOURNAL's Coöperating Editorial Board; Professors W. K. Estes, Indiana University; Francis W. Irwin, University of Pennsylvania; George A. Miller, Harvard University; Leo Postman, University of California at Berkeley; W. C. H. Prentice, Swarthmore College; and T. A. Ryan, Cornell University. We welcome them to the Board, to the inner circle in which the Editors of the JOURNAL have long shared the responsibility of accepting and rejecting manuscripts.

We have also to announce that Professor George L. Kreezer has again assumed the editorship of the Book Review Department which he carried so successfully during World War II and had to lay down when he was called into service. We bespeak for him your hearty coöperation. A review department is in the hands of its reviewers; it is what they make it.

K. M. D.

M. E. B.

E. B. N.

Harry Levi Hollingworth: 1880-1956

Harry Levi Hollingworth, Professor Emeritus of Psychology in Barnard College, Columbia University, died at his home in Montrose, New York, on September 17, 1956, at the age of seventy-six years. He had spent his entire academic life in that institution, from 1907 to his retirement in 1946, when he was named Professor Emeritus.

Professor Hollingworth was born on May 26, 1880, in De Witt, Nebraska, a village of some five hundred people, where the books accessible to him were limited mainly to the Bible and the Methodist Hymnal. He began his schooling at the age of eight years and shortly thereafter started working to help in the support of himself and his family. During the course of the next ten or twelve years he served, among other occupations, as church janitor, grocery store clerk, grave digger, carpenter, country-school teacher, church organist, and part-time preacher. When he was thirteen he was 'converted' and before reaching the age of sixteen years, was given an "Exhorter's License" in the Methodist Church. This enabled him to preach occasionally in rural churches. Indeed, he felt that he had received a divine call and decided to dedicate his life to the ministry and missionary work.

At about this time, however, a change occurred. He happened to see a mail-order catalogue containing a list of books selling at seventeen to forty cents each. He bought, as his meager funds permitted, volumes on Aristotle, Plato, Darwin, Spencer, and others. This reading completely changed his outlook on life and led him to seek a higher education and a career in philosophy and psychology.

Concerning this choice he wrote in later years:

So far as I can tell, this predilection represented a native taste or constitutional predisposition and was not due to anything that happened to me or that was taught me. I seem to have been born that way and have persisted in it in spite of obstacles.¹

He was admitted to the Preparatory Academy of Nebraska Wesleyan University at the age of twenty-one years with twenty-five dollars in his pocket, representing his sole savings from a couple of years of teaching and summer work. His ability and interest were soon recognized, and he was given special permission to attend an advanced college course in the History of Philosophy. Two years later, in 1903, he entered the University of Nebraska, where he worked under T. L. Bolton in Psychology, E. L.

¹ The above quotation and others that follow are taken from unpublished autobiographical material.

Hinman in Philosophy, H. B. Ward in Biology, and E. A. Ross in Sociology. Bolton made him assistant in experimental psychology in his second term, a post which he held during the next several years. Of this period in his training he later wrote:

So saturated did I become with philosophical issues and antinomies that, much to the disgust of my more empirical psychological colleagues, I have carried that interest with me into my own work.

He was graduated from the University of Nebraska in 1906 with membership in Sigma Xi and Phi Beta Kappa.

Columbia University accepted Hollingworth as a graduate student in 1907, and at the same time he was appointed an assistant in the Department of Psychology. His dissertation on "The inaccuracy of movement"² raised certain issues in his mind, issues which became the basis for the systematic point of view that dominated all of his later work.

Regarding this study he later wrote:

On the basis of my analysis of the experimental results, I concluded that there is no one-to-one relationship between inner sensory data and the external events of which these are in perception the signs. For each feature of a movement, such as its duration, speed, extent, force, I concluded, specific cues are established through experience. Movements are judged equal that have been *learned to be* equal, with no question of equality in the sensory patterns by which we know them. These may be equivalent but not necessarily equal. Furthermore, I realized, the sensory pattern is always a *context*, not an 'elementary sensation.' The correct judgment of our own movements involves the synthesis or comprehension into a single pattern of this context. Verdicts touched off by details not thus contextually checked are likely to be wrong. In psychophysics this leads to error; in perception, it leads to illusion; in clinical behavior it is the picture of the neurosis.

Thus two significant principles emerge. One is the use of partial sensory details as signs for situations first experienced more completely. The other is the factor of scope, that is, openness to the total context of signs presented by a given situation. In both of these there is emphasis on *cue* or symbol, and on *context* or pattern, insistence on a multiple cue theory, and rejection of a one-to-one doctrine of perception.

Koffka once commented to Hollingworth that this study had been one of the cornerstones of Gestalt Psychology.

After receiving the Ph.D. degree at Columbia in 1909, Hollingworth was appointed Instructor in Psychology at Barnard College. The year before, he had married Leta Stetter, and he found the salary of an instructor inadequate to cover his living expenses and the graduate education of his young wife. To increase his income he began to lecture and to conduct field studies and research in business psychology. These activities seemed to occupy the major portion of his time for several years. He had, however, no genuine interest in applied psychology as such and said in retrospect:

² *Op. cit.*, *Arch. Psychol.*, 2, 1909 (No. 13), 1-37.

My activity in the field of applied psychology was mere pot-boiling activity, and now that it is over there is no reason why the truth should not be revealed. My real interest, now and always, has been in the purely theoretical and descriptive problems of my science, and the books, among the twenty I have written, of which I am proudest, are the more recent ones which nobody reads.

Three of his major applied investigations were reported in psychological monographs and journals. They were "The influence of caffeine on efficiency,"³ "The psychological influence of alcohol,"⁴ and "The psychodynamics of chewing."⁵

The first of these projects was financed by the Coca Cola Company as part of a defense this company presented against a federal suit. It was an experiment designed with all the controls that are customary in the psychological laboratory. Not only did it lead to important findings concerning the effects of drugs, but to many others relating to individual differences and conditions affecting efficiency. Furthermore this psychological experiment aroused great respect and admiration among the biologists, pharmacologists, and psychiatrists who were likewise engaged in the project and set new standards in applied research. The second and third investigations followed the patterns of the first and led to equally important applied and theoretical results.

An excursion into vocational psychology, begun with a course of lectures at Teachers College, and experiments in the Barnard College Laboratory brought forth two books: *Vocational Psychology* (1916); and *Vocational Psychology and Character Analysis* (1928) together with numerous articles in the journals.

Hollingworth as Captain in the U.S. Army in World War I (1917-18) was assigned to a reconstruction hospital at Plattsburg, New York, where treatments were being developed for shell-shock cases. Out of his experiences there he formulated a theory of functional neuroses "which involved a specific intellectual weakness (lack of scope or sagacity); a past distressing episode; and the present effectiveness of partial features of the old episode in reinstating the similar feature of distress or incapacity." A wealth of cases from the hospital records supported his theory and the whole was organized into a book, *The Psychology of Functional Neuroses* (1920). He later turned again to this field in a more generalized fashion in his *Abnormal Psychology* (1930).

During all of these excursions into applied fields, Hollingworth never lost interest in theory. In fact, he seemed able in every instance to extract

³ *Op. cit.*, *Arch. Psychol.*, 3, 1912 (No. 22), 1-166.

⁴ *Op. cit.*, *J. abnorm. Psychol.*, 18, 1923, 204-237, 311-333.

⁵ *Op. cit.*, *Arch. Psychol.*, 34, 1939 (No. 239), 1-90.

something that would contribute to his systematic thinking. On one occasion he remarked that he never considered any of his investigations to lie in the field of applied psychology. He thought of them as straightforward efforts to discover the nature of certain facts and relationships and said of them:

There is, therefore, no real split personality involved in the simultaneous entertainment of interests so apparently diverse as those of technology, experimentalism, and systematic theorization. All these interests may usefully supplement one another in the work of science.

There was indeed a more immediate attention to systematic problems during these years, which manifested itself in numerous publications. Notable among these were his *The Psychology of Thought* (1926) and his address as President of the American Psychological Association on "Sensuous determinants of psychological attitude."⁶ In the former he made his first attempt at a formal statement of his systematic views. Of it he said:

The writing of this book gave me genuine pleasure and I am still fond of it. No one else appears to be. I have yet to see any reference to it in psychological literature.

A more extensive presentation of his theories is given in his *Psychology: Its Facts and Principles* (1928). This book was written as a general textbook in psychology and was so used for a number of years at Barnard College. It was not used in many other institutions because it was too difficult for an undergraduate without the sympathetic guiding hand of its author. *Educational Psychology* (1933) represented an attempt to apply his principles to the field of education and is the most convenient introduction to his systematic point of view.

The three fundamental principles upon which this systematic structure is based may be very briefly stated as follows: (1) The effectiveness of partial clues in reinstating consequents appropriate to larger antecedent contexts is the essential thing about mental phenomena and processes. The 'power of the symbol' lies here and this is what is meant by the spiritual. (2) The degree of scope—the openness of the individual to the present environment which is bearing down on him—as a characteristic in which organisms differ one from another is an equally fundamental fact and a heavy determinant of what is called the 'effectiveness' of mental processes. (3) All motives are distresses, that is, positive pangs, aches, hurts, and discomforts, and all action is in the endeavor to alleviate or reduce these initial distresses. Instead of conduct being the 'pursuit of pleasure,' it is a struggle away from goads.

⁶ *Op. cit.*, *Psychol. Rev.*, 35, 1928, 93-117.

Hollingworth published twenty-five books and monographs, seventy-five articles and papers, and thirty-six book reviews. As this survey shows, many of these were written in search of a systematic point of view. But systems come and go, and many remain the private possessions of their authors. Hollingworth knew this.

What are other evidences of his contribution to his chosen field? There were three honors that he prized highly: his election to the Presidency of the American Psychological Association in 1927; the award of the LL.D. in 1937 by his Alma Mater, the University of Nebraska; and the dedication of the Hollingworth Psychological Laboratories, Barnard College, in 1954. As a professor in an undergraduate woman's college, his influence is not easy to evaluate. Throughout his graduate life, however, he served as a member of the Columbia graduate seminar and regularly gave courses for the graduate students or made his undergraduate courses available to them. In addition he frequently guided research projects for doctoral candidates. At least fifty of his undergraduate students later received their Master's degrees in psychology, and twenty earned their doctorates. In the rapidly changing staff of the Department at Barnard, he had the privilege of guiding more than thirty young people into psychological careers.

Professor Hollingworth was a modest man, 'a lone wolf' he called himself, heartily disliking conferences and committees and working at his best alone. He would like, I think, to have this appraisal of him and of his work end with his own evaluation.

My life has been tame and mostly civilized. I have consorted with no kings, conquered no continents, tracked down no ferocious beasts, patented no inventions, and announced no startling discoveries. I have lived and labored; loved and lost; and through it all pursued an eager quest for a better understanding of that amazing thing called mind.

Columbia University

A. T. POFFENBERGER

BOOK REVIEWS

Edited by GEORGE L. KREEZER, Washington University

Behavior Theory and Conditioning. By KENNETH W. SPENCE. New Haven, Yale University Press, 1956. Pp. vii, 262.

The first of these Silliman Memorial Lectures stands apart from the remaining six and need not long detain the reader. It deals in superficial and rather biased fashion with the development of psychology as science, a topic considered appropriate on the ground that psychology has never before been represented in the Silliman series. Only in the second lecture does the author settle down to the principal and highly specialized business—far too specialized for an “audience composed of laymen, scientists from other fields, and psychologists of all varieties”—of ringing changes (at times engaging, often discordant, ever fragmentary) on traditional Hullian themes.

The learning treatise begins with a functional analysis of commonly employed experimental situations: selective learning (e.g. maze or Thorndikian problem-box, instrumental conditioning (e.g. runway or Skinner-box), and classical conditioning (e.g. Pavlovian salivary technique, or leg-flexion training with avoidable or unavoidable shock). The examples are puzzling until the basis of classification is grasped. Selective learning situations are the most complex. They involve hierarchies of competing stimulus-response connections (S_1-R_1 versus S_1-R_2 . . . versus S_n-R_n)—the competition is not, it should be noted, among responses elicited by a common stimulus (S_1-R_1 versus S_1-R_2 . . . versus S_1-R_n), as Hull and others before him conceived it. Any measured increase in the (initially low) probability of some differentially reinforced response, such as R_n , therefore reflects the strengthening, not only of S_n-R_n , but of its supporting receptor-orienting tendencies, along with the weakening of a host of other connections. Instrumental conditioning is purer; when the maze is reduced to a runway, and the orange-crate gives way to a stark, steel enclosure, R_n is much more likely to occur at the outset, because S_n is more “obtrusive,” and because the opportunity for competing responses is minimized. (The differences are regarded as quantitative rather than qualitative.) In classical conditioning, the greatest purity is achieved; the problem of receptor-orientation disappears entirely, and E is armed with a stimulus which (due to genetic factors or to previous training) directly elicits the response to be learned. One is startled to find avoidance training classified—first explicitly, and later, in the use made of a curve obtained by Brogden, implicitly—as a special case of classical conditioning in which the pattern of reinforcement changes from consistent to intermittent. Has the contingency of shock on failure of response no behavioral significance? And why is it that certain R_s , like running, are more stably conditioned by the avoidance technique, while others, like squeaking, are more stably conditioned by the non-avoidance technique? But perhaps the most insular feature of the treatment lies in the insistence that, when a neutral stimulus

is paired with brief, inevitable shock, the termination of the shock constitutes the "rewarding event" said to be essential for conditioning.

The author turns next (in the third lecture) to a review of some data on the shape of conditioning curves, from which he concludes that an initial phase of positive acceleration is typical (exceptions are noted, as when latency or magnitude of response is learned as such), and that level of motivation interacts with number of training trials to determine level of performance. These relations are derived (in the fourth lecture) from the equation $\dot{E}_L = [D \times H] - L - I_0$, where $\dot{E}_L$ is momentary excitatory potential, D is drive, H is habit-strength, L is the reaction-threshold, and I_0 is an oscillating inhibitory component. H is assumed to grow as a negatively accelerated, exponential function of the number of training trials, while I_0 accounts for the initial positive acceleration of the learning curve. Another inhibitory factor I_n , plays no part in the derivation of the curve because, unlike its Hullian forebears, it is unrelated to work, but develops as a (linear) function of "successive nonreinforced trials." Although reinforcement is regarded as essential for improvement of performance, no analysis is made of the manner in which it operates. The main enterprise here is curve-fitting, and the assumptions involved are of necessity too broad and too arbitrary at the present time to be of any great interest.

The author returns to functional problems in the fifth lecture, on "instrumental reward conditioning." To account for data on magnitude and delay of reinforcement in such situations, he uses the constructs K and I_t . The former, affected by frequency, amount, and delay of reinforcement, interacts multiplicatively with H (as in Hull) but additively with D —excitatory potential is a function of $H[D + K]$. The value of K is determined by the "vigor" of the "classically conditioned" fractional goal-response, r_g , which is regarded as its underlying "mechanism." (The theory thus is tied to a consummatory interpretation of reinforcement in spite of considerable contrary evidence.) The effect of a difference in amount of reward must in some cases be attributed to the conditioning of different r_g s, varying in vigor as a function of the unconditioned consummatory responses of which they are fractions, and in other cases to the different habit-strengths of a common r_g , which varies in vigor as a function of habit-strength. (Spence implies that the magnitude of the sudden shifts in level of instrumental activity which have been found to follow sudden shifts in amount of reward can be accounted for in these terms, but he attempts no formal deduction, and none seems possible.) Delay of reward, too, is reflected in K . To the extent that the learning situation changes with delay, the stimuli which elicit the response to be learned become less similar to those which are conditioned to r_g , and K is reduced by stimulus-generalization. In a situation unchanged by delay, r_g simply is extinguished. Delay also provides opportunity for the strengthening of competing stimulus-response connections, an outcome expressed by subtracting the quantity I_t from $H[D \times K]$.

Having decided (with Hull) that H is unaffected by magnitude and delay of reinforcement, the author (unlike Hull) considers that he has joined the "contiguity camp," at least insofar as "instrumental reward conditioning" is concerned. (This is a surprising development, since earlier in the same lecture the possibility of a "crucial experiment on the reinforcement-contiguity issue" is denied on the ground that it is "impossible to conduct an experiment without some form of motivation

and reinforcing event.") Reluctance to renounce the necessity of reinforcement for classical conditioning leads then to the tentative statement of a kind of inverted "two-factor" position—"that reinforcement (e.g., need reduction) plays a decisive role in the acquisition of the habit strength of classical conditioned responses but not in the case of instrumental conditioned responses"—the implications of which do not seem to have been thought through very clearly.

It should be noted, to begin with, that the contiguity treatment of the instrumental H does not involve as fundamental a reorientation to the problem as might first appear, since reinforcement still is necessary to forestall the growth of I_n . From the Hullian viewpoint, I_R is generated on all trials while sH_R increases only on reinforced trials. For Spence, H increases on all trials while I_n is generated only on unreinforced trials. In both cases, therefore, reinforcement increases reaction potential, and non-reinforcement reduces it, quite independently of K (the conditioned r_θ). The new position does not even avoid the awkwardness of a variable affected by reinforcement but unaffected by the amount or delay of reinforcement; it merely shifts the locus of the difficulty with the locus of the presumed operation of reinforcement—from H to I_n .

The two-factor formulation carries the implication that classical and instrumental conditioning are diverse processes—why otherwise should they be governed by different laws?—yet the distinction between them offered in the second lecture is purely procedural. The earmark of classical training seems to be that the response to be conditioned can be elicited at will by E , if not by bringing an appropriate US to the animal, then by leading the animal to the US (as when r_θ is connected to food-box cues). Given a dog that has learned to flex a leg to the sound of a bell for food-reward, one may ask whether, if it has not already reached its maximum, the H for flexion will continue to increase on unreinforced trials. The answer to this question scarcely can depend on whether electrodes have been attached to the limb, by which means E can elicit flexion when he so desires, yet no more satisfactory basis is provided for distinguishing classical conditioning. It is not limited to autonomic responses, because skeletal responses are spoken of as classically conditioned. It is not limited to reflexes, because the connection of a learned escape-response to a neutral stimulus is classified as an instance of classical conditioning. Perhaps only responses to motivationally significant stimuli are to be spoken of as classically conditioned, but that would beg the question of reinforcement which underlies the search for a distinction, since either the onset of an appetitive US or the termination of a noxious US may be called upon to supply the requirement. Another possibility is to be found in Hull's definition of classical conditioning as the formation of connections *de novo*, but then the level beyond which H may increase without reinforcement would have to be specified.

It is difficult to understand why Spence does not simply generalize his "contiguity" formulation and thus spare himself the task of finding a basis on which to distinguish classical conditioning. (If our analysis of that formulation is correct, he would then be no less a thoroughgoing reinforcement theorist.) The H of an r_θ , like that of an instrumental response, might be held to grow on all trials and to be negated by increments of inhibition only on unreinforced trials. Indeed, the analysis of K in "instrumental reward learning" already carries the implication that amount of reinforcement is important, not because the conditioning of r_θ is a

unique process, but only insofar as it affects the *opportunity* for that conditioning or the magnitude of the response conditioned. To be sure, it is concluded in the sixth lecture—a rather diffuse account of some conflicting data on the role of motivation—that the effect of *US*-intensity on eyelid-conditioning is due, at least in part, to “differential drive reduction” involved in the termination of the *US*. The results which appear to require this conclusion may, however, be attributed to variation in *D*, based (in accordance with a scheme proposed later in the same lecture) on the magnitude of r_s , a conditioned emotional response also elicited by the *CS*. A unified “contiguity” theory of this sort would, of course, still not have much application in any but highly selected realms of data.

Spence repeatedly emphasizes the narrowness of the “boundary conditions” within which his theory is intended to apply, but another, less defensible kind of narrowness is manifested in his failure to deal with clearly relevant data which seem to provide insurmountable difficulties for the theory. Although here and there in the book passing reference is made to a troublesome fact, the uninitiated reader might never guess that the overall empirical fit of the theory is far less impressive than the lack of it. While wondering at the amount of attention given to relatively insignificant matters, we must, of course, accept the implication that the treatment of such “important problems” as “partial reinforcement” was precluded only by “limitations of time.” Where critical issues *are* touched upon, however, the narrowness of outlook is unmistakable.

Consider, for example, the design of an experiment with the runway which seems to bring the model into line with early data on latent learning. After several unrewarded runs down an alleyway (let us call it A_x) to a similar goal-box (B_x), rats are put directly into B_x and fed. On the assumption that the r_s conditioned to the goal-box will generalize to the alley, an increase in the speed of running is anticipated. The validity of this interpretation is tested with another group which is given unreinforced runs in $A_x B_x$ but fed in B_y , a goal-box quite unlike A_x and B_x . In this case there should be less generalization of r_s and a smaller increase in speed of running. The proposed test is trivial, however, because the cognitive orientation which gave rise to the problem in the first place suggests precisely the same outcome, yet there are simple variations in the design which would yield unique deductions. For example, where one group of animals is pre-trained in $A_x B_x$ and a second group is pre-trained in $A_x B_y$, Spence must predict that (1) if both groups are fed in B_x , their subsequent performance will be equal; (2) if both groups are fed in B_y , their subsequent performance will be equal; (3) if either group is fed in B_x and the other in B_y , the subsequent performance of the group fed in B_x will be superior. Furthermore (4), where two groups are trained in $A_x B_y$, feeding in B_x will produce better performance than feeding in B_y . Deductions of this sort would be more worthy of the reader's patience.

In the seventh (and last) lecture, Spence works out some of the implications of his model for the relation of drive-level, frequency of reward, and delay of reward, to performance in a two-choice selective learning situation (S_A - R_A versus S_B - R_B). The apparatus employed presents two levers (S_A and S_B) far enough apart that “only one can be approached and operated at a time” (R_A or R_B), yet simultaneously visible from the choice-point, “thus evoking excitatory potentials to the two responses at the same moment.” The response which must occur, then, is that

which "has the greater excitatory potential at the moment." Although, under these conditions, r_g often must play a critical role in selection, the analysis of its operation is strikingly inadequate. If an instrumental tendency, such as S_A-R_A , is considered in isolation, one may think of its excitatory potential (E_A) as heightened by an r_g to S_A in the same manner as it is heightened by D . The conception changes, however, when, with S_A and S_B presented simultaneously, an r_g elicited by S_A is said to increase E_A but not E_B . How is the energizing effect of r_g channelized? The difficulty becomes particularly clear when both S_A and S_B tend to elicit r_g , but with different vigor, as when R_A and R_B both are rewarded, but with different amounts of food—how then are the fruits of the activity of a common consummatory mechanism correspondingly apportioned between the two excitatory potentials? Clearly, at least insofar as K is concerned, the question of response-selection has been begged.

What shall be said in conclusion about the book as a whole? Despite certain novel features and some interesting implications for experimental design, it is a thing, not of tomorrow, but of yesterday. It will intrigue specialists in learning, not in its own right, but as the work of its author—loyal and dauntless servant of a once-powerful, now tired tradition, painfully passing to its reward.

The Institute for Advanced Study

M. E. BITTERMAN

Clinical Studies of Personality. Edited by ARTHUR BURTON and ROBERT E. HARRIS. New York, Harper & Brothers, 1955. Pp. xiii, 836.

In some ways this is a book to be enjoyed. The first study in it, by Solomon Machover, is written beautifully, with compassion and genuine understanding. The character developed in it would be suitable for a novel. There appears to be true artistic insight by the writer. One must not be fooled, of course, by Machover's final formulation of the case, which constitutes, perhaps, a dutiful bow to the East, that is, in the direction of Freud, although it is not without literary merit: "The complaint of heartburn, of tendency to peptic ulcer, of asthma, and of precordial pain and frequent tachycardia and palpitations point to quite different psychological dispositions, yet each is appropriately expressive of the dynamics of Ezra's contradictory motivations, the asthma of his cry in desperation for the help of his rejecting mother, the gastric symptoms of his ambition's denial of his dependency, and the cardiac symptoms of his dependency's denial of his hostility. Perhaps it is the availability of alternate modes of suffering which prevents any one of them from becoming too serious" (p. 48).

Other contributors to this collection of case-studies make their bows to the East (although in most cases less dutifully) and display artistic insight (although often failing) to meet the literary standard set in the first study. For example: "From this marked 'oral fixation,' we would infer that Mr. W.'s aggression would tend to be expressed in spitting out of invectives, making 'biting remarks,' or 'chewing people out'" (p. 315).

Thus, the standard set in the first contribution is not always maintained. Most of the book consists of more or less dreary accounts of symptoms of suffering people and of attempts to treat them, with more or less fancy formulations based on more or less uncritical interpretations of projective and other test results. Here and there, a merciful break in this pattern occurs, as in Frank Barron's study

of a non-conformist, with its somewhat surprising denouement, and with its enticing title ("A Case Study of a Residual"). The longest study in the book displays perhaps the most unbridled speculation and calls to mind a finding of Holtzman and Sells in a study of the validity of the predictions of clinical psychologists—the more confident the judges, the less likely were their judgments to be valid.

Briefly, then, some parts of this collection can be enjoyed as literature. Further, some teachers will find in it useful illustrative material, and clinical students can discover in it a wide range of methods utilized by clinicians which is about all the editors seem to have intended. One might leave this review at that were it not for Henry A. Murray's introduction, reproduced in part from the initial effort of the editors (*Case Histories in Clinical and Abnormal Psychology*, 1947) and the reflections it stimulates about the current state of clinical psychology.

Murray's essay is essentially a passionate call—justified in terms of medical tradition—for idiographic studies of personality. He writes: "Hippocrates the Great, we know, was master of the case history method; and ever since his day, though for two thousand years no physician surpassed him in the use of it, this method has been the medical man's preferred way of accumulating knowledge. The adoption of the traditional procedure by the psychologist is simply the result of bringing him within the sphere of the psychiatric branch of medicine" (p. 13). Murray regards the case-history method as the preferred method for accumulating knowledge, although it is not precisely clear what kind of knowledge he has in mind. Murray does not deal with the profound epistemological problems his position involves. It is perhaps too elementary to note that the case-history method can produce, and sometimes has produced, interesting hypotheses, but little more, and that it is only one of a great many methods for the invention of hypotheses. Indeed, we have no science of science. There is no *one strategy* for the accumulation of knowledge, but the *tactics* are well mapped. They involve all of the experimental and statistical paraphernalia commonly included under the rubric of scientific method. Case-histories can be used to illustrate general laws of behavior arrived at through the tactic of experimentation, and they may be useful for many in the formulation of new hypotheses, but they cannot be taken as evidence for functional relationships in the same sense as the results of well-designed experiments.

Viewed from the standpoint of the possible additions to knowledge which may ensue from the reading of this book, the results are discouraging. Malmö's study of a case of hysterical deafness, previously reported elsewhere, is suggestive of a fruitful attack on a problem, as are one or two others. The general context, however, consists either of the type of speculation noted above or rather stereotyped recitals of psychometric data, all of which reflects accurately enough the current state of clinical psychology (one cannot take issue with the editors on this score) but does not provide much support for Murray's point of view. No new well-formulated hypothesis emerges from any of the contributions; nor is there anywhere in the book (we must exclude here the final chapter, entitled "Hemerography of Marry Ennis," which involves a new technique and is obviously part of a much larger undertaking beyond the scope of this review) a serious attempt to test alternative hypotheses against the case-data. (It might be interesting to see the results of such attempts and perhaps the editors might encourage them in their future collections.)

Rarely does an author refer to some principle of general psychology in the present work. It is difficult to foresee the accumulation of new knowledge from such productions.

If clinical psychologists are not induced by experiments like that of Holtzman and Sells, previously mentioned, and by writings like those of Meehl to examine the bases of their judgments, their present working hypotheses, and their methods they will soon find themselves entirely outside the realm of science—not because of their subject-matter, but because of the rejection in many of their claims and methods of the canons of scientific method. *Clinical Studies of Personality*, despite Murray's wholly devoted wishes, hopes, and expectations, does not reveal, either explicitly or implicitly, the path to knowledge in clinical psychology. It is questionable whether any benefit will come of the efforts of clinical psychologists, however heroic, until there has been a good deal of education concerning the different functions of science, applied science, and art.

University of Pennsylvania

JULIUS WISHNER

Nebraska Symposium on Motivation: 1955. Edited by MARSHALL R. JONES. Lincoln, University of Nebraska Press, 1955. Pp. x, 274.

The publication of the 1955 Symposium brings the third volume of a series. As has been characteristic of the previous symposia, the contributions to the present volume vary not only in the manner in which motivation is conceptualized but also with regard to the empirical data with which each contributor is concerned. The diversity of interest poses a difficult problem for the reviewer who must restrict himself to comments upon only a few of the many issues that are raised. In the reviewer's opinion, this diversity tends to reduce the value of the symposium. Although each paper is commented on by at least two of the participants, more effort should have been devoted to the clarification and resolution of conceptual differences. While reports of research are both informative and stimulating, the symposium would be of greater value if the papers were limited to an analysis of theory and research in some specific problem-area, such as the acquisition of motives, substitute or vicarious satisfaction, and the effects of punishment, to mention but a few possible topics.

The first participant, Maslow, reasserts the theme of his recent book, *Motivation and Personality*, that there is a class of "growth" motivations which are characteristic of psychologically healthy individuals and which, by virtue of their different antecedents and different consequents, must be distinguished from "deficiency" motivations. Although Maslow's arguments are both plausible and appealing, the definition of a psychologically "healthy" individual and the measurement of the elusive growth motives still remain as thorny and critical problems. Dealing with quite different phenomena, Olds reviews in detail his significant experiments on the "reinforcing" effects of electrical stimulation of the various structures comprising the rhinencephalon, or limbic system. He proposes the admittedly speculative but intriguing hypothesis that low-level neural firing in certain limbic structures activates behavior while high-level neural firing arrests or inhibits it.

Young, in the context of his hedonic theory, reports a group of interesting experiments dealing with the development and modification of preferential behavior

in rats. Scattered throughout his paper are a number of strong criticisms of Hullian theory, despite the fact that none of the reported experiments is crucial to or was designed to test that theory. While an explanation of Young's data in terms of drive-reduction theory may be clumsy, we can not agree with Young that the theory has been shown to be "completely inadequate." Dealing with human preferences, Helen Peak presents a disparity-theory of motivation, and systematically relates the constructs of attitude and motive. A theoretical advantage of conceptualizing the discrepancy between two psychological processes as the stimulus to action is the ease with which cognitive sources of motivated activity can be handled. In addition, the disparity concept, within the tradition of homeostatic approaches to motivation, was specifically constructed to account for the characteristic "persistence" of motivated behavior. The essence of this construction is to shift the referent of motivation from the desire for a particular goal to the discrepancy between the goal and some other psychological state, *e.g.* the individual's current status with respect to that goal. The value of this approach to motivation will depend upon the specification of the pertinent dimensions along which disparities may vary, the specification of the conditions under which a particular disparity will lead to action, and the development of techniques for measuring disparities which would enable one to predict the intensity, persistence, and direction of motivated behavior.

McClelland's paper, in a quite different vein from the other contributions, is devoted to an exploration of the social consequences of achievement-motivation—in particular, to the effect of the Protestant Reformation upon the growth of capitalism. The evidence that he provides linking independence training to Protestant values, achievement motivation to independence training, and Protestantism to economic development, indicates that Protestant values are consistent with and reinforce certain economic values and behavior. To assume the priority of Protestantism, however, is gratuitous, and does not fit the facts any better than assuming the priority of economic and class changes.

To buttress his argument regarding the relationship between achievement-motivation and economic progress, McClelland reports a significant association between high *n*-Achievement and vocational interests of a financial nature. This suggestion of a similarity between the profit motive and the achievement motive raises a serious question regarding the generality of *n*-Achievement which apparently no longer pertains to any striving to attain a standard of excellence. Furthermore, while the goal-objects of *n*-Achievement may have been reduced in number, the antecedents of *n*-Achievement will probably prove to be more numerous and complex than McClelland postulates. "What makes Sammy run" may not be similar to the motivations of Jay Gould or Benjamin Franklin. Intense achievement-strivings may be based upon sibling rivalry and feelings of inferiority as well as upon the high value of achievement-symbols stemming from independence training.

Rotter, the last participant in the symposium, outlines the essentials of his social learning theory while stressing the importance of situational factors in the prediction of behavior. His theory has certain formal similarities to that of Hull, although it deviates from the latter in a number of important respects. In Rotter's modification of Hull's formula for reaction potential, *reinforcement value* and *expectancy* are substituted for *drive* and *habit strength*, respectively. This formula-

tion, while more explicit than that of Peak, for example, is less parsimonious and would have greater difficulty in handling such problems as the activation and cessation of behavior.

While the differences in point of view among the participants are numerous, a few areas of agreement can be discerned. The drive-reduction statement of the law of effect is decidedly unpopular. Olds and Rotter commit themselves to an empirical law of effect. The dissatisfaction of McClelland, Young, and Maslow with the drive-reduction hypothesis stems from their contention, supported by their own research as well as the work of Sheffield and Harlow, that increments as well as decrements in stimulation can function as reinforcers. Maslow's statement that Olds's data "clearly kill need-reduction theories of motivation forever—or at least the monistic ones" is overly optimistic. Olds himself draws no such inference and, in fact, is quite careful to point out that his results may be a consequence of numerous effects including a cessation of activity in the limbic structure resulting from electrical stimulation. A number of the contributors emphasize situational determinants of motivation and are in accord with the increasing tendency to externalize motives. Perhaps the presumed shift in our culture from inner-directed to other-behavior is being reflected in our scientific theories.

University of Pennsylvania

SEYMOUR FESHBACH

The Psychology of Human Differences. By LEONA E. TYLER. Second edition. New York, Appleton-Century-Crofts, 1956. Pp. viii, 562.

In a plethora of psychological textbooks, differential psychology can boast of only two systematic expositions: Anastasi's volume of 1937 (revised in 1949 with the aid of Foley) and Tyler's *opus*. Compared to the first (1947) edition, the present revision has grown in width (by half an inch), in height (by one inch), and in thickness (from 420 to 562 pp.). As before, only studies in which psychological characteristics are assessed in quantitative terms are considered, with synthesis of research findings as the goal, but there have been important changes in organization and emphasis, e.g. the space devoted to elementary statistical concepts is substantially reduced.

Considerations of measurement in psychology, together with a historical introduction, form Part One. It is unfortunate that distributions which are symmetrical (not skewed) but which clearly differ in kurtosis continue to be labelled "normal" (p. 26), in spite of the expressed statement (p. 22) pointing out that such a designation is incorrect "unless a mathematical test has been made to see whether the relationships between x and y . . . satisfies a certain mathematical equation." While the shaping of the new science is considered only briefly in Chapter One, William Stern's important role appears to merit the citation of his mature work *Die differentielle Psychologie* (1911, 3rd ed. 1921).

Part Two is devoted to individual differences in various aspects of man's behavior (intelligence, scholastic achievement, vocational aptitude, personality, interest, attitude, and perception). Whereas in 1947 perception was discussed on less than one page, in the present edition a whole chapter is devoted to it. The author feels that research on perceptual differences occupies a strategic position in the contemporary differential psychology, that the results obtained thus far are encourag-

ing, and that research on perceptual variables may contribute importantly to the study of abilities and personality as well as of social attitudes.

Part Three deals with the varieties of group differences. It would perhaps have been more satisfactory to separate the material into two parts, one dealing with bio-social groupings (on the basis of sex, race, nationality, social class, and age), and the other concerned with extremely deviant individuals. To the considerations of the mentally deficient and the unusually gifted, falling into the second category, one could profitably add the psychological portrait of the mentally ill, even though some overlap with the chapter on individual differences in personality might be a problem, and strict adherence to the quantitative viewpoint would reduce drastically the volume of relevant literature.

Part Four is labelled "Factors Producing Differences." The single chapters deal with the relationship between physical and mental characteristics (where no direct, one-way etiology is necessarily involved, and where the new developments in the study of bodily composition have not been considered), influence of the environment, and heredity. The concluding chapter bears a fitting title—"Toward a Science of Human Differences: Achievements and Challenges."

In this work the author does not limit herself to a juxtaposition of different views, but strives for synthesis. Emphasis is placed on the *directions* in which the current research is moving and on developments opening new vistas. Not all has been said, not all experiments completed. In reference to the recent work on individual differences in perception it is noted that the "textbooks of the future may be organized in terms of basic variables only dimly sensed at the present stage of research" (p. 501).

On the research level, the interaction between the hereditary potential and the effects of education and experience is stressed as one of the most challenging problems, especially as regards the *patterns* of traits which represent the very essence of human individuality. While the concept emphasized most strongly by the author is that of *statistical significance*, we deal here with matters of acknowledged and profound *practical importance* for the scientific management of all societies in which the welfare of concrete human beings is the criterion of the soundness of educational, legal, economic, and even medical policies and institutions.

The author envisaged intelligent upper-division and graduate students of psychology as her principal readers. They are being served a meal that is appetizing as well as nutritious. At the same time, at least some parts of the volume will be found useful by a variety of persons who are faced daily with the facts of individual differences, whether it be in education, social work, or a multiplicity of administrative tasks in business, industry, government, and the armed forces. In the days when the eyes of the world are focussed on America's handling of its racial problems, one wishes that scientific facts on racial differences (and similarities), carefully summarized and interpreted by Tyler, might be given wide recognition and serve as a basis for an intelligent social action.

Last, but not least, individual and group differences are not limited to behavior. They are clearly present in bodily size, shape, and composition, in biochemical characteristics and physiological functions. Specialists in these fields would do well to familiarize themselves with the psychologist's sophisticated methodology. Such a familiarity, with awareness of the possibilities and limitations of the present-day

methods for multivariate analysis, would enhance the formation of a comprehensive, multidisciplinary Biology of Human Differences.

University of Minnesota

JOSEF BROŽEK

The Individual Psychology of Alfred Adler. A Systematic Presentation in Selections from His Writings. Edited and annotated by HEINZ L. ANSBACHER and ROWENA R. ANSBACHER. New York, Basic Books, 1956. Pp. xxiii, 503.

Students of personality theory have long bemoaned the fact that Adler never produced a single, systematic, readable work representative of his position, as did Freud, and, to a lesser extent, Jung. For the most part, Adler's work is known to the American public through his semi-popular writings and through summary statements in secondary sources. Much of the original German material never has been translated. As a result, there has been a tendency to undervalue his work, which can be understood, since summaries and popularizations always tend to oversimplify. The Ansbachers have set out to correct this state of affairs. They have produced a volume which brings together Adler's views on personality and development, the problems of abnormal psychology, and the problems of allied areas, such as criminology and social psychology. Their purpose is not only to present a systematic picture of Individual Psychology, in the fashion of a text, but also to relate this work to the larger psychological context.

An initial chapter, written by the editors, constitutes the major effort toward fulfilling the last-mentioned aim. In it, a clear differentiation between the views of Freud and Adler is attempted. Freud is seen as an objective psychologist while Adler comes off with the subjective designation as defined in the work of Karl Jaspers. Adler's position is related to those of Stern, Wertheimer, Lewin, Dilthey, Vaihinger, Rogers, and Heidegger. Freud's work is considered in relation to that of Pavlov, Thorndike, Hull, Dollard and Miller, and so forth. A point of contention arises with regard to the modern analytic views represented by Horney, Fromm, and Sullivan. While these writers have been referred to as neo-Freudians, the Ansbachers reiterate what is becoming the popular view that they might more accurately be called neo-Adlerians. A highlight of this chapter is the statement of twelve clear, concise, fundamental propositions of Individual Psychology as gleaned by the editors from Adler's writings.

The book has two major sections. One deals with personality theory and development, the other with the more practical aspects and applications of Individual Psychology.

In the first section one finds all of the concepts that are readily identified as Adlerian, such as masculine protest, fictionalism, striving for superiority, social interest, and style of life. These concepts serve as headings for chapters which contain selections portraying their importance in the system. Sometimes a concept is traced historically to show changes in Adler's thinking. At other times the various applications of a particular concept may be brought into focus. As in other sections of the book, the selections are introduced by editorial remarks which serve to orient the reader.

Most of the second section is devoted to Adler's views on neurosis and other behavior-disorders, including chapters on the neurotic disposition, symptoms and defenses, dreams and early recollections, and the problems of treatment. The last

few chapters strongly reflect Adler's very marked practical outlook. They are concerned with the deviant personality, as evidenced through crime, addictions, and perversions, and also with the general life-problems, which include occupational choice, marriage, adolescence, sexuality, and old age. A final chapter deals with problems in social psychology. Here one can find Adler expressing himself on religion, prejudice, the social responsibility of the psychologist, Bolshevism, and similar topics from the standpoint of Individual Psychology.

The editors by design make no attempt to evaluate Adler's views. Their effort seems rather to be directed towards a better appreciation of his role in the history of psychology. Numerous instances are pointed out where early Adlerian concepts later appear in the Freudian system, the aim of the editors being perhaps to counteract the popular notion that Adler's primary contributions are nothing more than reactions to psychoanalytic doctrine, and that Adler's claim to fame, so to speak, is based on his dissent from Freud. One receives the impression of a vigorous and creative individual who could hardly be held in check by the imposition of another's views. Adler is portrayed as an independent wave in the mainstream of holistic, field-oriented, subjective positions in psychology.

This book is an important addition to the literature of personality and clinical psychology. One will recognize in it much that has already been assimilated in professional psychology as well as in the general population. (Many popular child-rearing practices of the last twenty years will be identified as emanations from Adler's pronouncements.) Its most general value, however, will be as a much needed source book for Individual Psychology. There are, of course, the usual difficulties of continuity, repetitions, and flow, common to most books of this kind, but they are held to a minimum by a thoroughly professional editorial job. The Ansbachers are to be congratulated. They have, by their wise selection, excellent organization, and insightful comments, enriched Adler's work.

Princeton University

IRVING E. ALEXANDER

Language, Thought, and Reality: Selected writings of Benjamin Lee Whorf. Edited by JOHN B. CARROLL. Cambridge, The Technology Press of Massachusetts Institute of Technology; New York, John Wiley & Sons; and London, Chapman & Hall; 1956. Pp. xi, 278.

"A change in language can transform our appreciation of the Cosmos." This was the theme to which Benjamin Lee Whorf (1897-1941) lent his talented pen during the 1930s, which stimulated widespread discussions among anthropologists, linguists, philosophers and psychologists during the 1940s, and which seems to arouse cautious qualification and pessimistic reservation in the 1950s. "Concepts of 'time' and 'space,'" said Whorf, "are not given in substantially the same form by experience to all men but depend upon the nature of the language or languages through the use of which they have been developed."

At one level of discourse, and especially when one has fallen under the charm of Whorf's writing, the thesis seems obvious, almost banal. It must be true that an appropriate notation is the first step toward understanding a problem (evidence from mathematics is convincing here), and it also must be true that some natural languages provide better systems of notation for some problems than do others (the variety of the world's languages makes this point highly probable). Yet at another level, when one tries to draw hardheaded, empirical consequences from

the thesis, the whole argument seems so vague as to be more confusing than helpful. In order to compare the way two languages appreciate the Cosmos, one gets involved in difficult technical problems of translation, in the murky depths of the psychology of thought and understanding, in metaphysical questions about the true nature of the Cosmos, and in the age-old problem of the meaning of meaning. Surely there must be limits to the extent to which changes in language can transform our world-view, for we are not completely free to take any view we like if we want to survive, but where these limits must be drawn is hard to say—the discussion degenerates into a quarrel about the relative importance of the symbol and the referent in the processes of cognition. Because of this combination of obviousness and vagueness, Whorf's views will probably continue to stimulate discussion and research—it is hard to escape the feeling that there is something important here if only we could formulate it more precisely.

John Carroll has done a masterful job of editing Whorf's papers, some of which are published here for the first time. His introduction combines an excellent biography with an appreciative statement of the central points in Whorf's linguistic philosophy. Fortunately, the enthusiastic foreword by Stuart Chase is quite short.

Psychologists interested in the related problems of language, thought, and cognition will be grateful to Carroll for making Whorf's major works available in a single volume. One can only wish he had done so a decade earlier.

Harvard University

GEORGE A. MILLER

Psychology of Exceptional Children. Edited by WILLIAM CRUICKSHANK. Englewood Cliffs, Prentice-Hall, 1955. Pp. xiii, 575.

This symposium, said to be "devoted solely to the . . . influence of physical deviation upon the normative growth and development of children and young people" (p. vi), actually is broader in scope. Of the eleven chapters, one is on assessment (with minimal regard for physical deviations), one is on speech (which includes non-physical influences), one is on the gifted (with negligible reference to the incidence of superiority among the physically deviate), one is on psychotherapy (much needed, but little exploited), and one is on mental retardation (not usually considered as physical deviation, and not here emphasizing physical fitness).

I like this book, which recognizes the need for collation of rapidly accumulating information about 'special' children, but I am a bit puzzled about its potential audience. Perhaps it is intended for students in the psychology of physical and mental deviation. It is frankly research-oriented, and has the familiar academic flavor, as though the authors were reluctant to talk about deviates as people (the current trend) or to reveal intimate acquaintance with them. One feels that the contributors' wealth of personal clinical experience might well have made a more human outlook profitable. Such case-materials as are included are more formal than personal, more scientific than professional.

The bibliographies are more selective than representative, with minimal use made of work prior to about 1940. (The references are hard to appraise because of their concealment among the notes appended to the chapters.) In the urge for modernity, many excellent 'old' papers are ignored, with serious loss to a research-oriented volume. Are historical developments of so little moment?

I find here the customary inequalities of edited volumes—some chapters are better than others—and I miss chapters on behavior disorders, CNS impairment, and

other types of exceptionality implied in the title, but the book lends itself readily to endorsement without carping.

Bellingham, Washington

EDGAR A. DOLL

The Meaning of Nationalism. By LOUIS L. SNYDER. New Brunswick, Rutgers University Press, 1954. Pp. xv, 208.

Professor Hans Kohn, who is noted for his numerous studies on nationalism, remarks in a foreword to this book that nationalism is a phenomenon of great complexity and that the coöperation of many disciplines is required in order to achieve an understanding of it. Louis L. Snyder, Professor of History at the College of the City of New York, has made the first attempt at such an interdisciplinary study by drawing on sources not only in history but also in sociology, economics, political science, anthropology, psychology, and psychiatry, including psychoanalysis. The result is a clear, though detailed and extended, appraisal of the many forms of behavior and states of mind to which the term *nationalism* has been applied. In stating the aim of the book Snyder expresses the hope that it will "provide a kind of extensive definition such as might have been presented by the editors of Webster's unabridged dictionary if they had had several hundred pages at their disposal instead of a few inches of space." The book is more, however, than a mere dictionary-definition of a single term. Rather, it is a description of the many forms and expressions which nationalism has taken.

Snyder begins with the semantic problem and then, after analyzing the concepts of nation and nationality, he discusses the various suggestions for classifying forms of nationality, explores the relevant economic factors, and distinguishes between patriotism and nationalism. Finally he considers the problem of national character and the spurious concept of a national soul. He arrives at the general conclusion that the two major meanings of nationalism are concrete (geographical) and ideal (psychological). He insists that nationalism "should be considered first and foremost a state of mind, an act of consciousness, a psychological fact." As an expression of the mind it appears in a "good" form in which it is free from an attitude of superiority, and in an "evil" form in which it is a divisive force, generating fears, frustrations, and aggressions. Its function is to be that of a "socially approved symbol used by modern society in its search for security." It is interesting to a psychologist that nationalism is thus looked upon primarily as a psychological problem. From here it should not be far to a recognition that psychology is a basic science for the study of history.

The question might be raised whether it would not be wise and scientifically more appropriate for professional workers to use a word like *nationalism* most sparingly, or not at all, and instead deal with a specific social phenomenon without encumbering the discussion with the many other meanings of the term. In other words, would we not gain in precision and clarity by focusing upon this or that specific referent rather than asking "What does the term nationalism mean?" This is not said in order to detract from the undoubted value of this book but rather to point out that one of its significant functions might be to help a student of this problem to isolate the various referents of the term in question and to indicate with which one he is concerned.

University of Colorado

KARL F. MUENZINGER

Scientific Method in Psychology. By CLARENCE W. BROWN, and EDWIN E. CHISELLI. New York, McGraw-Hill, 1955. Pp. ix, 368.

The appearance of a comprehensive book on this subject seems long overdue. It is not uncommon for undergraduates, after being exposed to that traditional first chapter on 'Psychology as Science' in an introductory text, to accumulate a good many credits toward the B.A. (or even to graduate) without ever coming to grips with the logic, methods, and problems involved in psychological research. The student too often gets lost in the minutia of data-collection, 'doing statistics,' and report-writing in courses on experimental psychology courses, and mastery of the typical research-laden text frequently leaves the student with a mass of specific findings but little integrated knowledge of methodology. This book should do much to help fill that void in undergraduate training.

The presentation is divided into three parts. In the first of these, "Some General Concepts about the Scientific Method," the authors have done a rather unique job of presenting a basic orientation to science without losing the scientist (*i.e.* the psychologist) in the shuffle. He appears as a person with special training and skills which lead him to operate in ways that are different from those of the non-scientist; but he emerges as a human being instead of a flamboyant genius or an antiseptic automaton of the laboratory. In this reviewer's opinion, this emphasis is more than just scientifically wholesome. It seems especially appropriate to the orientation and motivation of the student of psychology.

The introductory statements for Part II, "Steps in the Scientific Method," provide a good example of the authors' style and outlook: "The scientific method does not consist of a single set of steps followed in some invariable chronological order. To a large extent, the chronology of a scientific experiment is tied to the ups and downs of the scientist's motivations. At some time in the early training of a student of science, however, there is need for a systematic treatment of those steps that are fundamental to any scientific attack upon a problem. In presenting such a treatment, no claim should be made that the steps are well delineated one from another or that the order of the steps is established regardless of the nature of the scientist's progress." For this part of the book, the authors have selected a major research problem to illustrate the concepts of each chapter. For example, highway-accident prevention is used to illustrate the problems of Chapter 7, "The Definition and Delimitation of a Scientific Problem." While some of the other chapters use several problems or actual research-findings illustratively, the use of a single problem within a chapter provides unusually good continuity for the ideas presented.

The third part of the book, "Some Individual Scientific Procedures," addresses itself to the following topics: stimulus-variables, the subject, and procedures for quantifying behavior. The problems associated with each are amply illustrated. The final chapter gives four examples of the application of the scientific method to field studies.

The authors report having used the material presented here in mimeographed form for a wide variety of college courses. Like their *Personnel and Industrial Psychology*, this book is clearly and interestingly written, with appeal to a broad range of students. There are times, however, when the materials seem to be somewhat pedantically 'over-organized' or slow-moving for graduate purposes. In general, the book provides

sufficiently important content to justify its use as the text for a course on psychological methodology for upper-division majors.

San Francisco State College

ROBERT E. DREHER

Research Films in Biology, Anthropology, Psychology, and Medicine. By ANTHONY R. MICHAELIS. New York, Academic Press, 1955. Pp. xvi, 490.

This volume was written in the stimulating context of the Scientific Film Association of Great Britain, an organization without an exact American counterpart, which is broadly concerned with the use of photography in scientific research and teaching. The knowledge and skill of the author are extensive, and it is obvious that a great deal of effort has gone into his book. The treatment is scholarly—complete with historical orientation, theoretical discussion, and quantitative data—and detailed. For example, after describing with great thoroughness and clarity the equipment-arrangments and techniques in biology, the author reviews applications to such specific problems as tissue-sectioning, bacterial growth, virus research, mitosis, cancer studies, and animal behavior. The treatment of other areas is equally systematic and thorough. Investigators who have made no extensive use of cinematography will find many new ideas.

The volume is well illustrated. Equipment-systems are diagrammed or photographed in detail sufficient for duplication, and the results of cinematography are shown as well as possible by still photographs. There is a bibliography of 1,490 titles and a convenient subject index.

Pennsylvania State University

C. R. CARPENTER

Mental Disorders in Later Life. Edited by OSCAR J. KAPLAN. Second edition. Stanford, Stanford University Press, 1956. Pp. ix, 508.

This book, first published in 1945, now appears in a second edition which includes four additional chapters: Benjamin Malzburg's "A Statistical Review of Mental Disorders in Later Life," Franz J. Kallmann's "Genetic Aspects of Mental Disorders in Later Life," Clive M. McCay's "Food for the Later Years," and Curtis T. Prout, Edward B. Allen, and Donald M. Hamilton's "The Uses of Electric Shock Therapy in Older Patients." Almost all books which are composed of chapters written by different authors are marked by a certain unevenness. The difficulty is accentuated in the present volume by the fact that the contributors come from such widely divergent fields as genetics and nutrition. Some of the chapters (e.g. "The Presenile Dementias") are medical presentations of specific conditions, while others are rather technical treatises dealing primarily with methodology (e.g. "Psychological Aspects of Mental Disorders in Later Life").

The chapters by Kallmann, by Shock, and by Lewis are particularly illuminating. Kallmann draws together the scattered literature relevant to the role of genetics in longevity and in the mental disorders of later life. His own investigations, which have been somewhat scattered in the literature, are correlated with the findings of others, and a firm basis is established for an evaluation of findings, both old and new, relevant to the length of life. Shock's chapter, on "Physiological Aspects of Mental Disorders in Later Life," carries with it a bibliography of 342 items which, according to the author, is not comprehensive. The chapter presents a clear summary of present day knowledge relevant to the

topic, but even more valuable is its accenting of the points on which information is not available. Certain statements in other chapters, particularly those dealing with clinical medicine, would have had to be modified had the authors had the opportunity to have read Shock's chapter before writing theirs. All too much of the writing on the topic of mental hygiene is composed of pious platitudes and well-worn, but unsubstantiated advice. Nolan Lewis' chapter, on "Mental Hygiene in Later Maturity," might well serve as a model for future writing in the field of mental hygiene. His observations are keen, his advice realistic, his statements are factual, and his tolerance is exceptional.

This volume is recommended. One wishes that more of the areas on which psychology borders or with which it overlaps had compilations on the same high level of competence.

Columbia University

CARNEY LANDIS

Personal Influence: The Part Played by People in the Flow of Mass Communications. By ELIHU KATZ and PAUL F. LAZARSFELD. Glencoe, The Free Press, 1955. Pp. xx, 400.

The first part of this book examines in detail the role of interpersonal relations in the mass-communication process. The traditional orientation of the mass-persuasion studies is criticized for neglecting interpersonal relations which function to *reinforce* and to *relay* mass-communications. To delineate the reinforcement function, many of the group-dynamics studies on uniformities (conformity) is reviewed. This analysis is penetrating and comprehensive. Next, the practical implications (*i.e.* for change in attitude and behavior) of group-dynamics theory are discussed, considering the group as a "medium" and as a "target" of change. In the latter category, the authors present a definitive critique of the famous group-decision studies, offering novel proposals for theoretical and experimental clarification. The relay function of interpersonal relations is then described through an analysis of those group-experiments in which the quantity, content, and direction of communication constitute the dependent variable. Specifically considered are the effects upon these communication-processes of sociometric structure, status-structure, social "climate," common interest, and leadership. The analysis of the different bases of leadership ("nominational," social location, and cultural certification) and the different functions of leadership (origin, sanction, and diffusion of ideas) is original and stimulating. In all, the authors achieve their objective of exposing the enormously complex interpersonal processes that intervene between the mass-media and their audience. The arguments will convince many of the futility in regarding mass-persuasion as a problem of individual psychology.

The second part of the volume reports on a sample survey of 800 women from Decatur, Illinois, a "representative" city in terms of several demographic and economic indices. In essence, an attempt was made by repeated interviews of the same respondents to describe actual influences and their sources in the areas of marketing, fashions, opinions about public affairs, and movie-going. What is the relative effectiveness of personal contact and the mass media in each of these areas? What are influential people like—in terms of their age-marital status, their socio-economic status, their gregariousness? Such are the types of question answered by the data. It turns out, for instance, that movie-going is mainly influenced by young girls.

Matters of public affairs, on the other hand, are mainly influenced by higher status "opinion leaders."

It is obvious that such findings are of great practical moment. It is not so certain what specific significance they have for current or for new theories of social influence. Although the data show clearly that the concrete lines of influence—who influences whom—will vary with the content of the influence, the two well known theories underlying influence "pressures" reviewed earlier in the book—"membership motivation" and "social reality"—do not appear to require any kind of modification or differentiation as a result of this finding. On the other hand, the idea that certain roles—opinion leaders—function to mediate the influence-attempts of other groups and the mass-media seems to point the way to a more general social psychology—one that will embrace intergroup as well as intragroup relations.

The book, in sum, will broaden the purview of social psychologists. Further, what the authors have to say about social influence will interest and inform psychologists concerned with individual motivation and learning.

University of Pennsylvania

ALBERT PEPITONE

The Analysis of Fantasy: The Thematic Apperception Technique in the Study of Personality. By WILLIAM E. HENRY. New York, John Wiley and Sons, 1956. Pp. xiii, 305.

The title of this well-written treatise is somewhat of a misnomer, since the author concerns himself primarily with the products of imagination as elicited by Murray's *TAT* (in some cases modified for use with Indian Ss). Conspicuously lacking is a treatment of the structure and function of fantasy as one of the 'higher mental processes.' The subtitle more apt.

The first half of the book offers a clearcut description of the process of *TAT* interpretation developed and employed by Henry. The second half illustrates the use of Henry's interpretative system, first by comparisons of stories evoked by specific pictures, and then by four detailed analyses. The final chapter contains a useful account of the social-stimulus properties of the standard pictures. Two appendices complete the book: the first contains some concrete pointers on test-administration; the second is a comprehensive bibliography on the *TAT*.

Early in the book, the author—unlike most writers in this field—explicitly states the major dimensions of his theory of personality. In an exposition remarkably free from the hoopla often encountered in writings on projective tests, Henry sketches a theory of personality organization which does not neglect social and cultural features. For example, he explicitly proposes that "The stories are not exclusively a projection of inner states. They are rather a resultant of these inner states, in response to a personal definition of generalized social situations, as represented by the pictures" (p. 35).

To the reviewer, Chapter 6, which provides "a conceptual framework for individual case analysis," is a high point in the book. Every psychologist who employs clinical instruments does, of course, develop his own framework. This one, however, has much in its favor. If it had been developed and widely adopted ten years ago, then today, instead of hodge-podge, we might have accumulated a body of data which would have permitted systematic and comparative studies of persons and of groups. In this brief review, suffice it to say that the framework is

comprehensive, realistic, and readily communicable. The illustrative analyses are convincingly impressive. They are patent applications of Henry's social psychological theory. Portraying the Ss as true-to-life people, in places they approach the representational literary form.

Applicable to the general field of projective testing as well as to *The Analysis of Fantasy* are two related criticisms: (1) the assumption of validity of interpretations and (2) the failure to slant the interpretations toward the prediction of events. Henry acknowledges that a systematic treatment of validity is lacking. In this connection, one might assert that rigorous studies of validity can be planned and prosecuted now that an easily communicable set of concepts is available. For the prediction of specific behavior, the *TAT* must be used with extreme caution, warns Henry. "To describe the personality of a subject is an easier task than describing the projection of that personality into behavior in its social context" (p. 36). The emphasis on personality-description over behavior-prediction is a shortcoming that should be noted.

Although not a manual in the usual sense, the interested reader will find many tips, hints, and suggestions that will be helpful in interesting *TAT* stories. *The Analysis of Fantasy* should be regarded as required reading by all students who intend to use the *TAT* or similar methods in personality research or in clinical practice. Forthright, clear, and modest in its claims, this book is the best introduction to thematic apperception now available.

University of California

THEODORE R. SARBIN

Industrial Psychology and its Social Foundations. By MILTON L. BLUM. Harper and Brothers, New York, 1956. Pp. xvi, 612.

Here is a revised edition of a book which first appeared in 1949. It differs from the first edition more in organization than in content. Its twenty-one chapters, which are fundamentally the same chapters as in the earlier book, are grouped into six parts. Part I describes the jobs of the industrial psychologists and exemplifies their methodology with a detailed account of the Hawthorne study; part II considers worker motivation, job satisfaction, morale, the techniques of attitude measurement, the special problems of unemployment; part III discusses leadership and industrial warfare; part IV includes discussions of personnel selection, placement and training, the working environment, accidents, fatigue, and time and motion studies; part V considers the problems of consumer research, advertising, and selling; part VI is a one-chapter summary of the book. Each chapter concludes with a summary and bibliography. There is a 12-page appendix containing a brief description of the psychological tests most commonly used in industry.

The book is full of content and is organized concisely enough that the businessman who reads it should find it a valuable reference. The conclusions from studies of various problems are set forth as guides to action. Only once does Dr. Blum slip from his position of basing conclusions on scientific research. This slip occurs when he advises on "How Not to Buy" using anecdotal data. The illustrations do not always do their job of illuminating the text, and (occasionally) even confuse because of inadequate legends.

The author contends that the "promotion of more harmonious employer-employee relations in a democratic society demands that these relations have an extended

democratic base, that the social aspects of these relations be understood and that the basis for mutual trust be rooted in statements of goals and the search for scientific knowledge." After reading this statement in the Preface, the reviewer was surprised to find the text devoted primarily to a presentation of the results of research and disappointed not to find more emphasis upon methodology and a greater development of the arguments supporting the social origins of interpersonal relations. The book suggested by the statement in the Preface would be an exciting book to read. The students and the executives, union leaders, owners of small businesses, and employees, for whom this book was written, should find it interesting, helpful, and informative, but they probably will not develop as great an understanding of the methodology and social foundations of industrial psychology as the author hoped they would.

Knox College

ROBERT S. HARPER

THE AMERICAN JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXX

JUNE, 1957

No. 2

PROBABILITY, LEARNING, THE STATISTICAL STRUCTURE OF CONCEPTS, AND THE MEASUREMENT OF MEANING

CHARLES M. SOLLEY and SAMUEL J. MESSICK, The Menninger Foundation

Although there have been many discussions about the meaning of 'meaning,' there have been few attempts to measure meaning.¹ This paucity of measurement stems not only from difficulties as to how to measure, but also from uncertainty as to what to measure. The possibility should be considered that several different kinds of meaning may exist for one concept and that various measuring devices may be reflecting different aspects of meaning. When different aspects of a single concept exist, they are probably related to the different ways and, in some cases, to the different situations in which individuals learn that concept. Thus, a consideration of concept-formation might have important implications for the systematic description and measurement of various aspects of meaning.

This experiment is an attempt to relate the measurement of meaning to certain hypotheses concerning the formation of meaning in a statistical framework. It considers the meaning of a concept to be a reflection of probability-relationships among and within the referents of the concept. For example, consider one way in which we could learn a concept such as 'Italians.' First we might experience a large number of referents to which the label 'Italians' is applied. A specific Italian might be tall, happy, skinny, and quick; another might be short, happy, fat, and quick.

* Received for publication November 14, 1955. This research was done at the University of Illinois.

¹ C. E. Osgood, The nature and measurement of meaning, *Psychol. Bull.*, 49, 1952, 192-237. For more general discussions see S. I. Hayakawa, *Language in Action*, 1941, 214-261; Charles Morris, *Signs, Language, and Behavior*, 1946, 1-31; C. K. Ogden and I. A. Richards, *The Meaning of Meaning*, 1923, 185-208.

Whenever S experiences a specific Italian, he experiences (or samples) a combination of perceptual events.² Each combination of the potential perceptual events has a certain probability of occurring in the population of referents to which a common label is being applied. Somehow, as the individual gets more and more experience with more and more referents, communalities are abstracted, and a concept emerges—a process exemplified by the learning of Hull's Chinese characters.³ Several abstractions can be made, however, for any given population of referents, and communalities may be combined in different ways for each abstraction. It seems, then, that if we knew more about probability-relationships among the referents of a concept, we could systematically construct what might be called the 'statistical structure' of the concept; *i.e.* we could describe the concept's meaning in terms of the likelihood that certain combinations of the various aspects will occur.

These different combinations of the aspects of a concept may be reflected differentially by various instruments designed to measure meaning. For example, the 'semantic differential' may measure only one statistical aspect of the meaning of a concept; whereas other techniques, such as word association or the method of triads, might reflect different statistical aspects of the same concept.⁴ If we knew more about the statistical structure of concepts, then, it might be possible to specify some relationships between that statistical structure and the various devices used to measure meaning.

STATISTICAL STRUCTURE OF CONCEPTS AND PROBABILITY-LEARNING

The following hypothetical example illustrates what is meant here by the statistical structure of the perceptual characteristics of referents. Let us consider an imaginary tribe of 24 South Pacific natives that happens to have the distribution of perceptual characteristics shown in Table I. Fifteen of these tribesmen are tall and nine are short; eighteen are happy and six are sad. All 15 of the tall tribesmen are happy and none is sad, while only three of the short tribesmen are happy and six are sad. If an explorer (E) were to land on this island and were to meet these natives one at a time, he would gradually acquire information concerning this combination of characteristics; *i.e.* he would learn that most of the tribesmen are tall and happy, none of the tall tribesmen is sad, and short tribesmen are likely to be sad. If we were to ask E to guess what the characteristics would be for the next native he was about to meet, we could gain some

² W. K. Estes, Toward a statistical theory of learning, *Psychol. Rev.*, 57, 1950, 94-107; W. K. Estes and C. J. Burke, A theory of stimulus variability in learning, *Psychol. Rev.*, 60, 1953, 276-286.

³ C. L. Hull, Quantitative aspects of the evolution of concepts, *Psychol. Monogr.*, 38, 1920 (No. 123), 1-86.

⁴ Osgood, *op. cit.*, 222-232; W. S. Torgerson, Multidimensional scaling: I. Theory and method, *Psychometrika*, 17, 1952, 401-419.

information about the way in which he learned these attributes; *i.e.* if *E* were to guess beforehand whether each native in turn was going to be tall or short and happy or sad, we could plot a learning curve for each of the four combinations of attributes. Recent studies on probability-learning suggest the result that, as *E*

TABLE I
PERCEPTUAL CHARACTERISTICS OF A HYPOTHETICAL TRIBE

Characteristics	Happy	Sad	Total
Tall	15	0	15
Short	3	6	9
Total	18	6	24

TABLE II
FREQUENCY OF OCCURRENCE OF VARIOUS COMBINATIONS OF CHARACTERISTICS FOR
FOUR HYPOTHETICAL TRIBES OF STICKMEN

Characteristics	Tribes			
	A	B	C	D
tall-skinny-happy-white	12	3	0	0
tall-fat-happy-white	0	4	0	0
tall-skinny-sad-white	3	8	0	0
tall-fat-sad-white	0	0	0	0
short-skinny-happy-white	0	5	0	0
short-fat-happy-white	0	0	0	0
short-skinny-sad-white	1	0	0	0
short-fat-sad-white	4	0	0	0
tall-skinny-happy-black	0	0	4	0
tall-fat-happy-black	0	0	1	0
tall-skinny-sad-black	0	0	0	5
tall-fat-sad-black	0	0	0	0
short-skinny-happy-black	0	0	3	8
short-fat-happy-black	0	0	0	4
short-skinny-sad-black	0	0	12	3
short-fat-sad-black	0	0		

becomes more and more familiar with these natives, he would begin guessing each combination of characteristics more and more in proportion to their 'true' frequency-distributions.⁵

This simple example describes how learning increments reflect the proportional occurrences of various combinations of perceived attributes, and it also suggests how this statistical structure might be investigated by an experiment in probability-learning.⁶

To study the probability-learning of combinations of perceptual characteristics, it was decided to construct an artificial concept for which the 'true' frequency-distribution of attributes would be known. The concept chosen was 'Tribe,' and

⁵Egon Brunswik, Probability as a determiner of rat behavior, *J. exp. Psychol.*, 25, 1939, 175-197; R. R. Bush and Frederick Mosteller, A mathematical model for simple learning, *Psychol. Rev.*, 58, 1951, 313-323; H. W. Hake and Ray Hyman, Perception of the statistical structure of a random series of binary symbols, *J. exp. Psychol.*, 45, 1953, 64-74.

⁶E. C. Tolman and Egon Brunswik, The organism and the causal texture of the environment, *Psychol. Rev.*, 42, 1935, 43-77.

the referents for this concept were a series of drawings of little men whose perceptual characteristics varied with respect to the following bi-polar attributes: *tall-short*, *happy-sad*, *fat-skinny*, and *black-white*. Drawings of 20 little men, who came to be known as 'stickmen,' constituted a 'tribe'; four tribes were constructed, each with a different distribution of perceptual characteristics. These distributions are given in Table II.

This table indicates that Tribes A and B are identical with respect to the marginal frequencies of occurrence for the different stimulus-characteristics; *i.e.* there are 15 tall and 5 short stickmen in each tribe, 12 happy and 8 sad, 16 skinny and 4 fat, and 20 white. These two tribes, however, are rather different with respect to the joint occurrence of specific combinations of perceptual characteristics; *e.g.* in Tribe A there are 12 tall, skinny, happy, white stickmen, while only three have that description in Tribe B. Thus, with respect to marginal characteristics, Tribes A and B have identical 'meanings,' but with respect to another statistical aspect of the referents, mainly the joint occurrence of combinations of characteristics, the tribes have somewhat different 'meanings.' Tribes C and D are mirror images of Tribes A and B, respectively; *e.g.* instead of having 12 tall, skinny, happy, white tribesmen as in Tribe A, Tribe C has 12 short, fat, sad, black ones.

So, if a measuring instrument should reflect only the marginal probability-relationships, Tribes A and B should be found to be identical with respect to measured meaning. Tribes C and D should also be identical in this aspect, although their 'meaning' would not be the same as for Tribes A and B. If, however, the measuring instrument reflects the joint occurrence of characteristics or some combination of joint occurrence plus the marginal characteristics, then the meanings of Tribes A, B, C, and D should be differentiated. Since the meanings of concepts, even in terms of only two semantic aspects, involve characteristics varying in more than one way, it seems reasonable to base semantic measurement upon multidimensional models. One such device for measuring meaning multidimensionally is described in the next section, and the adequacy of its measures is evaluated in terms of the statistical theory of concept-referents discussed above.

THE SEMANTIC DIFFERENTIAL

Mosier made one of the few attempts to measure the meaning of concepts with a unidimensional psychometric scaling procedure, but the complexity and probable multidimensionality of the meaning domain makes a one-dimensional approach seem inadequate for complete coverage.⁷ When Osgood and Suci developed the *semantic differential*, a multidimensional analysis of meaning-structure as a basis for more complete measurement became available.⁸ By factor analyzing the inter-correlations among 50 scales which had been identified by such bipolar adjectives as *good-bad*, *strong-weak*, etc., they first obtained some idea of the dimensionality of meaning.⁹ Three orthogonal factors were extracted: an evaluative dimension, a potency dimension, and an activity dimension. It was now reasonable to attempt to find some measure that would enable experimenters to locate concepts in such a

⁷ C. I. Mosier, A psychometric study of meaning, *J. soc. Psychol.*, 13, 1941, 123-140.

⁸ C. E. Osgood and G. J. Suci, A measure of relation determined by both mean differences and profile information, *Psychol. Bull.*, 49, 1952, 251-262.

⁹ Osgood and Suci, A factor analysis of meaning, *J. exp. Psychol.*, in press.

three-dimensional semantic space. To do this, Osgood required subjects to judge a concept's position along each dimension separately, having selected a small number of relatively pure-factor bipolar scales to represent each of the three semantic dimensions. In other words, a group of subjects (*Ss*) are asked to judge the position of a concept on a set of bipolar scales representing the evaluative dimension, then on a set representing the potency dimension, and finally on a set representing the activity dimension. The average scale-position of the concept on each of these dimensions locates it as a point in the semantic space.

If, however, the meaning of a concept depends upon probability relationships among characteristics of the concept's referents, the question now arises as to which statistical aspect of meaning is being reflected by the semantic differential. Or, put another way, for what statistical aspects of concepts is the semantic differential 'valid'? In responding to the semantic differential *S* is asked to check the position of a concept on a bipolar scale (for examples see Fig. 1). For a concept like 'tribe of stickmen' and a scale like *tall-short*, *S* would have to judge the degree to which stickmen were characterized by tallness or shortness. The judgment process required of *S* in this case might be conceptualized as the formation of a scale position for the concept in terms of the resultant of the two vectors composing the scale—the 'tall' vector vs. the 'short' vector. That is, *S* has experienced a number of 'tall' stickmen and a number of 'short' stickmen and he checks at the point describing the average of his experiences. If the scale consists of seven categories running from +3 (very tall) through zero to -3 (very short), and if equal intervals are assumed, then the resultant score (an expected value) for the concept 'tribe of stickmen' based *only* on the marginal stimulus-probabilities would be:

$$E_{\text{score}} = 3p_{\text{tall}} - 3p_{\text{short}}$$

where E_{score} is the expected score, p_{tall} refers to the proportion of times tall stickmen occur in the tribe, and p_{short} refers to the proportion of times short stickmen occur in the tribe.

This expected score is based only upon the marginal characteristics of the concept, so Tribes A and B would have identical resultant scores. If this simple two-vector model adequately represents the way in which an individual checks a scale on the semantic differential, then Tribes A and B should both be placed at the same predicted point. If Tribes A and B are both placed at the same point but not at the predicted one, then the semantic differential is still only reflecting the marginal characteristics of the concepts, since these are the only characteristics for which Tribes A and B are identical, but the two-vector score model is not an adequate representation of this. If Tribes A and B are placed at different points on the semantic differential scales, then not only is the two-vector model inadequate, but the instrument is reflecting something more than marginal meaning.

The use of this resultant score to describe an individual's performance on the semantic differential assumes that he has previously learned the perceptual characteristics of the referents. Also, if the measuring device is to have a chance to separate the meanings of Tribes A and B, the *Ss* must also learn the joint occurrences of the combinations as well as the marginal frequencies. So in this experiment the methods used enabled the *Ss* to become familiar with the referents; the details of the procedure are given in a later section.

Since marginal probabilities for the perceptual characteristics of the artificial

concepts of tribe are known, expected values can be computed for each of the scales *tall-short*, *fat-skinny*, *black-white*, and *happy-sad*, thus generating an expected profile of meaning for each of the four concepts of tribe. It must be emphasized here that these expected profiles of meaning are based only upon the marginal characteristics of the concepts, hence Tribes A and B have identical expected profiles, as do Tribes C and D. If, however, the combination frequencies affect the judgment process involved in responding to the semantic differential, some systematic deviation would be expected from the predictions based only on the marginals. Hence, identical profiles for Tribes A and B obtained from the semantic differential would indicate that the judgment process required by that instrument takes only marginal characteristics of concepts into account; whereas different profiles for these two tribes would indicate that the combinations of characteristics are being considered in the procedure. These statements also hold for Tribes C and D.

These predictions, of course, hold only on the assumption that the learned combinations are proportional to the combinations built into the concepts. Even if learning were not complete, however, the above predictions could still be tested by generating expected scores based upon final probabilities of response rather than requiring that the input probabilities be learned exactly.

METHOD

Subjects. Forty undergraduates were used as Ss. Ten were randomly assigned to each of the four 'tribes' or conditions for learning.

Materials. The materials used in the learning phase of this experiment are described in Table II. Each tribe was composed of 20 cards, upon each of which appeared a schematized figure with a certain combination of characteristics. Two sets of drawings were made for each tribe, and each set was arranged in random order.

The materials used in the measurement phase consisted of a semantic differential composed of 20 scales (as shown in Fig. 1). Four of the scales corresponded specifically to attributes of the stickmen, *i.e.* *tall-short*, *fat-skinny*, *happy-sad*, and *black-white*. The other 16 scales were included to obtain some additional information about possible semantic generalization for the tribes.

Procedure. (1) *Learning phase.* The procedure for the probability-learning was to have each S guess the characteristics of the figure on the top card of the deck, the deck being face down. For example, he would guess whether the drawing on the next card was tall or short, then whether it was happy or sad, then whether it was fat or skinny, and then whether it was black or white. Each S thus made one guess for each of the four bipolar attributes on each trial. The card was then turned over, allowing S to scan the picture for approximately 5 sec. The card was then removed from view and S made his four guesses about the next picture. When guesses had been made for all 20 pictures in this way, another random order of the same deck was presented. This procedure was continued until all the cards in the deck had been seen four times. In pilot work, it had been found that almost all of the learning had occurred by the time S had gone through the deck twice.

(2) *Measurement phase.* The experimenters (Es) gave each S a copy of a semantic differential composed of 20 scales (as shown in Fig. 1) immediately following the phase of probability-learning. S was told to rate on each of the scales the 'tribe of stickmen' he had just seen. The successive intervals were indicated on each scale, but S was permitted to check anywhere along the length of the scale rather than being forced to use restricted scale positions.

RESULTS

(1) *Learning phase.* Summaries of the results of the probability learning are given in Tables III and IV. These tables show that the frequencies

TABLE III

SUMMARY OF LEARNING OF STATISTICAL STRUCTURE OF TRIBES A AND B
(Ten Ss learned each Tribe. Scores are average frequency of guesses for blocks of 20 guesses, a 'trial unit'.)

Characteristics	Tribe A, Trial units					Tribe B, Trial units				
	1	2	3	4	Input	1	2	3	4	Input
white-tall-happy-skinny	4.9	10.0	9.7	10.8	12.0	3.6	6.1	5.5	6.8	3.0
white-tall-happy-fat	1.2	.3	.1	.0	.0	2.6	2.1	1.9	2.4	4.0
white-tall-sad-skinny	2.5	3.3	4.8	3.3	3.0	2.6	5.3	5.6	6.5	8.0
white-tall-sad-fat	3.3	.3	.0	.2	.0	1.4	1.5	1.4	.7	.0
white-short-happy-skinny	.9	.3	.5	.4	.0	2.5	1.9	3.5	2.0	5.0
white-short-happy-fat	1.6	1.5	.8	2.2	.0	1.6	1.0	.5	.7	.0
white-short-sad-skinny	.8	.2	.8	.5	1.0	1.2	1.6	1.3	.7	.0
white-short-sad-fat	3.3	3.5	3.1	2.6	4.0	.7	.5	.3	.1	.0
black-tall-happy-skinny	.2	.2	.0	.0	.0	.6	.0	.0	.0	.0
black-tall-happy-fat	.9	.1	.0	.0	.0	.2	.0	.0	.0	.0
black-tall-sad-skinny	.3	.2	.0	.0	.0	.7	.0	.0	.0	.0
black-tall-sad-fat	.5	.0	.0	.0	.0	.2	.0	.0	.0	.0
black-short-happy-skinny	.2	.0	.0	.0	.0	.1	.0	.0	.0	.0
black-short-happy-fat	.4	.0	.0	.0	.0	.9	.0	.0	.0	.0
black-short-sad-skinny	.4	.1	.0	.0	.0	.4	.0	.0	.1	.0
black-short-sad-fat	1.6	.0	.0	.0	.0	.7	.0	.0	.0	.0

TABLE IV

SUMMARY OF LEARNING OF STATISTICAL STRUCTURE OF TRIBES C AND D
(Ten Ss learned each Tribe. Scores are average frequency of guesses for blocks of 20 guesses, a 'Trial unit'.)

Characteristics	Tribe C, Trial units					Tribe D, Trial units				
	1	2	3	4	Input	1	2	3	4	Input
white-tall-happy-skinny	1.8	.3	.0	.2	.0	.8	.2	.1	.0	.0
white-tall-happy-fat	1.0	.0	.1	.0	.0	.6	.1	.1	.0	.0
white-tall-sad-skinny	1.9	.4	.2	.1	.0	1.1	.1	.0	.0	.0
white-tall-sad-fat	.6	.0	.1	.0	.0	.3	.1	.0	.0	.0
white-short-happy-skinny	.3	.2	.0	.1	.0	.3	.3	.1	.0	.0
white-short-happy-fat	.8	.0	.0	.1	.0	.3	.2	.2	.0	.0
white-short-sad-skinny	.4	.1	.1	.0	.0	.5	.0	.0	.0	.0
white-short-sad-fat	.5	.0	.0	.2	.0	.4	.1	.3	.0	.0
black-tall-happy-skinny	2.8	.8	1.7	1.8	4.0	.6	.4	.5	.6	.0
black-tall-happy-fat	.8	.5	.6	.9	1.0	2.8	2.8	2.8	1.9	.0
black-tall-sad-skinny	1.9	1.2	.8	.7	.0	1.3	.6	.8	.8	.0
black-tall-sad-fat	.8	1.5	.3	.2	.0	3.3	3.0	1.8	2.2	5.0
black-short-happy-skinny	.6	.4	.2	.1	.0	1.2	1.2	.9	1.1	.0
black-short-happy-fat	1.8	6.1	5.5	3.8	3.0	2.2	5.0	6.9	7.7	8.0
black-short-sad-skinny	.8	.4	.2	.7	.0	1.4	2.7	2.4	1.8	4.0
black-short-sad-fat	3.1	8.0	9.0	11.1	12.0	2.9	3.2	3.1	3.9	3.0

of the Ss' guesses converge, on the average, to the input frequencies for the combinations of schematized perceptual characteristics. This result indicates that the Ss learned the joint probabilities for the occurrence of

combinations of characteristics, from which the marginal probabilities could be generated, and not just the marginal aspects of the concepts.

An inspection of Tables III and IV also reveals that the Ss tend to over-guess the frequency of occurrence of *happy*. In only one tribe, Tribe D, were they, however, overguessing *happy* at the end of learning to any marked extent. Although there were more *sad* figures in Tribe D than *happy* ones, the Ss in this group were guessing in almost reversed proportions at the end of learning, *i.e.*, instead of guessing 12 *sad* and 8 *happy*, they were guessing 8.7 *sad* and 11.3 *happy*. This seems to indicate that there might well be perceptual biases existing prior to this kind of experiment and that a design counterbalancing the conditions, such as the one used in the present study, is to be recommended.

TABLE V
EXPECTED SCORES FOR TRIBES BASED ON INPUTS IN LEARNING, $E(S)$, AND ON FINAL RESPONSES IN LEARNING, $E(R)$, WITH EMPIRICALLY OBTAINED MEANS, $\bar{X}$.

Characteristic Scales	Tribe A			Tribe B			Tribe C			Tribe D		
	$E(S)$	$E(R)$	$\bar{X}$	$E(S)$	$E(R)$	$\bar{X}$	$E(S)$	$E(R)$	$\bar{X}$	$E(S)$	$E(R)$	$\bar{X}$
tall-short	1.5	1.2	2.2	1.5	1.8	2.0	-1.5	-1.7	-2.1	-1.5	-1.3	-1.6
happy-sad	.6	1.0	1.1	.6	.6	.5	-.6	-.6	-1.5	-.6	.5	.9
black-white	-3.0	-3.0	-3.0	-3.0	-3.0	-3.0	3.0	2.6	2.6	3.0	3.0	3.0
fat-skinny	-1.7	-1.5	-1.7	-1.7	-1.8	-1.6	1.7	1.8	2.2	1.7	1.7	1.6

Since the Ss had now learned certain specifiable aspects of the meaning of these concepts, *i.e.* the joint probabilities for the occurrence of combinations of characteristics as well as the marginal probabilities for the relevant dimensions operating in each case, techniques purported to measure meaning could now be applied to see which semantic aspects are adequately reflected by various measurement devices.

(2) *Measurement phase: (a) Comparison of scores.* Table V shows the expected scores (concept scale-positions) based on (i) frequencies of marginal input; (ii) frequencies of marginal responses, and (iii) the mean scale-positions empirically obtained. This table shows that the average discrepancy between the mean scores obtained and those predicted is about a third of a scale unit (on the seven-unit scale used by Osgood and Suci). This discrepancy is comparable to average error-discrepancies found in other investigations of the reliability of the semantic differential. Thus, since the predicted scale-positions for the concepts based only upon their marginal relationships were not markedly different from the scale-positions obtained from the semantic differential, it may be concluded that this instrument as it is currently used reflects only the marginal characteristics of a concept's meaning.

Since the predicted scores were based upon a simple two-vector model that assumed equality of intervals along the scales of the semantic differential, the similarity between the scores predicted and obtained can be taken as evidence that the scale-intervals are at least not markedly unequal. The results also show that the semantic differential is valid for at least one statistical aspect of the meaning of concepts, but, on the other hand, they also indicate that the instrument as it is currently used does not adequately reflect all aspects of meaning.

TABLE VI

SUMMARY OF DATA FOR THE SEMANTIC DIFFERENTIAL AND *F*-VALUES AMONG TRIBES FOR EACH SCALE

Characteristics Scales	Tribes (means)				<i>F</i> -values
	A	B	C	D	
tall-short	2.2	2.0	-2.1	-1.6	153.3†
happy-sad	1.1	.5	-1.5	.9	11.2†
black-white	-3.0	-3.0	2.6	3.0	—§
fat-skinny	-1.7	-1.6	2.2	1.6	94.1†
weak-strong	-.3	-.9	-.4	-.4	.4
large-small	1.2	.7	-.7	-.9	5.1†
sweet-sour	1.2	-.2	-.3	.7	3.6*
good-bad	1.1	.3	.2	1.2	1.6
clean-dirty	1.2	1.4	.3	.5	1.6
honest-dishonest	-.7	-.5	-.5	-1.1	.5
heavy-light	-.5	-1.1	1.1	1.6	8.4†
rugged-delicate	.0	1.2	.3	.9	1.7
sharp-dull	.6	.8	.2	-.5	1.6
beautiful-ugly	-.2	-.4	-.8	-.3	.4
sick-healthy	-1.2	-1.2	-.7	-.9	.3
relaxed-tense	.7	.5	-.2	1.5	1.7
hard-soft	.8	1.0	-.1	.1	.9
rich-poor	-.5	-.5	-.3	-1.0	.6
rough-smooth	-.9	.7	.3	.0	1.9
fast-slow	1.1	1.0	-1.4	-.8	10.1†

* 5% significance level.

† 1% significance level.

‡ 1/10% significance level.

§ No *F*-value was computable since Tribes A, B, and D had no within-group variance.

(b) *Analyses of variance.* In addition to the comparison of the scores, simple one by four analyses of variance were made among the concepts for each of the 20 scales. The means obtained for each of the concepts on each of the scales, with corresponding *F*-values, are shown in Table VI. As predicted, significant differences were obtained for each of the four input scales with no significant differences between either Tribes A and B or C and D—with the exception of Tribes C and D on the *happy-sad* scale. Even this deviation, however, is understandable in terms of the final level of learning for Tribe D. The output prediction based upon the final level of learning for Tribe D indicated more 'happy' than 'sad' stickmen,

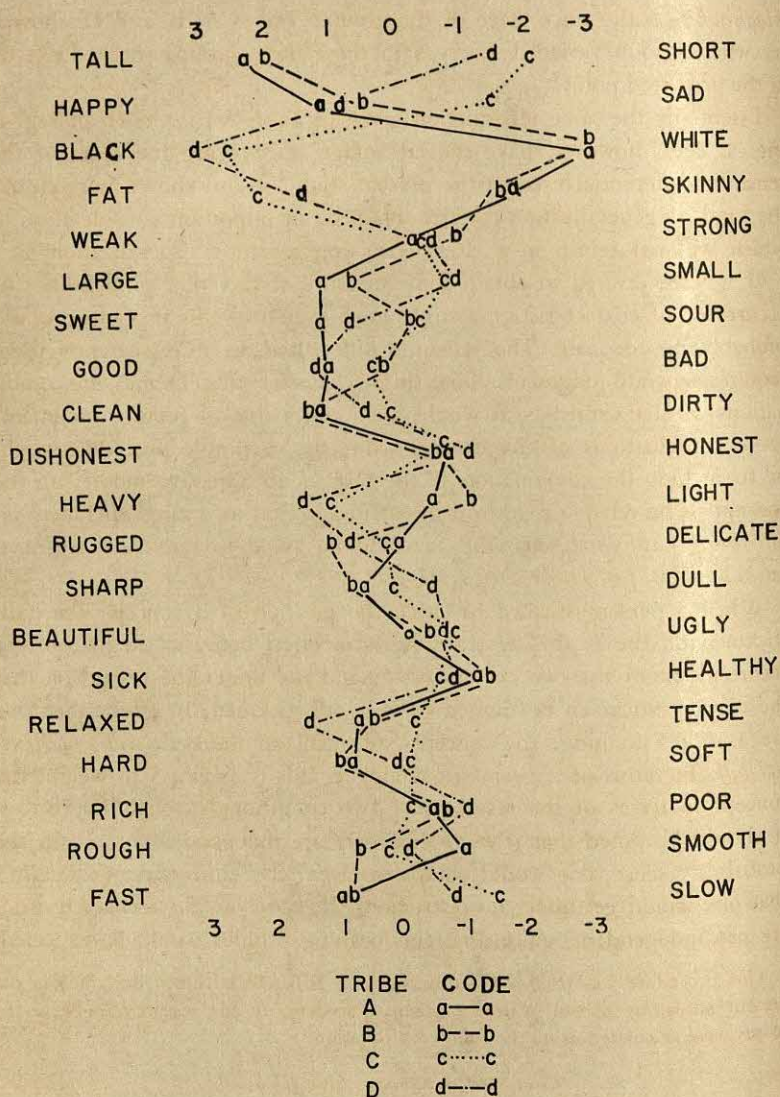


FIG. 1. MEAN PROFILES OF CONCEPTS OF FOUR TRIBES ON 20 SEMANTIC DIFFERENTIAL SCALES

although the reverse situation was true in terms of the input proportions. In addition, it should be noted that no *F*-value could be computed for the *black-white* scale, since three of the groups, Tribes A, B, and D, showed no within-group variance. Every *S* in these three groups checked exactly at the predicted point.

Essentially the same information given in Table VI is shown in Fig. 1. Fig. 1 does, however, have the advantage of showing the form of the semantic differential used in the present study; it also shows more clearly the kind of generalization-pattern obtained, an important consideration to which we will return in a later section on semantic generalization.

If it were desired to obtain some measure of the joint probabilities of occurrence of characteristics, a slight modification of this procedure would probably be adequate. The semantic differential, as it is currently used, requires separate judgments about single characteristics. To measure combinations of characteristics, it would seem reasonable to require judgments about combinations of characteristics directly. A simple procedure might be to include the combination in the task as an adjective modifying the concept to be rated; *i.e.* instead of, or in addition to, rating 'stickmen' on *tall-short*, *happy-sad*, etc., the *Ss* could be asked to rate 'tall' stickmen on *happy-sad*, *fat-skinny*, etc., scales.

When a person is asked to check the position of a concept like 'tall' stickmen on the *happy-sad* scale, he is in effect being asked to consider the resultant of the two vectors *happy* and *sad* under the restriction that the concept stickmen be limited to only tall-stickmen. In effect, then, we are asking *S* to judge the concept 'stickmen' on the scale *happy-tall* vs. *sad-tall*. In terms of a geometric analogue, this is asking *S* to weight the concept in terms of the resultant of two conditional probability vectors. It should be noted that if *happy* and *sad* are independent of *tall* in the probability sense, one would obtain from *tall* the same effect (resultant) that one would get under no restrictions. If, however, these characteristics are not independent, quite different meaning profiles would be expected.

This procedure was tried out, using the same tribes of stickmen described in the present study, by Sylvan Wiley, a graduate student. In his research the expected scores were calculated using the following equation:

$$E_{\text{score}} = 3p_{\text{tall}} | \text{happy} - 3p_{\text{short}} | \text{happy}$$

All possible combinations of conditional probability were calculated, and the above is only one specific example. The results of this research will not be reported in full in this paper but Fig. 2 illustrates some of the predictions and empirical results.

Fig. 2 shows that fairly accurate predictions were made for the semantic

differential under this modified procedure, using the judgmental model assumed to be the vector of conditional probability. Although these results are encouraging, the procedure must be tested further before generalizing too far. The results do indicate, however, that we may be able with a little manipulation, to reconstruct the joint frequencies of occurrence of stimulus-characteristics for the referents of a concept. That is, since $p(x,y) = [p(x,y)]/[p(y)]$, where $p(y) \neq 0$ and since we can estimate $p(x,y)$ and $p(y)$ from the scores on the semantic differential, we can thus estimate $p(x,y)$ by $p(x,y) = p(x,y) p(y)$.

(3) *Semantic generalization.* Since little is known about semantic generalization, 16 scales (other than those involved in the above predictions) were included to obtain some preliminary information about it. Although it is a crude test, the *F*-value for each of the other 16 scales was used as

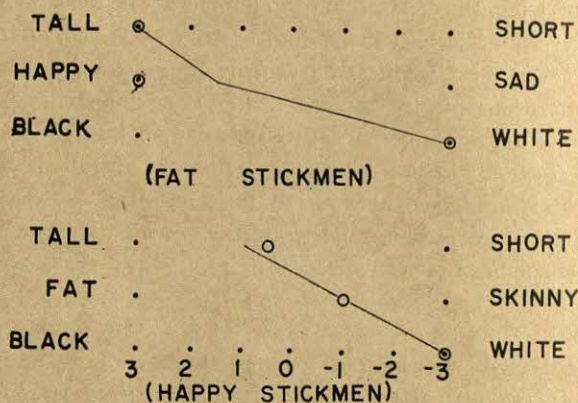


FIG. 2. PREDICTED PROFILES FOR 'FAT' AND 'HAPPY' STICKMEN OF TRIBE B WITH MEAN RATINGS ON EACH SCALE

an index of a significant amount of generalization. It might be argued that this statistic only indicates significant differentiation between the four groups on the scales for which the *F*-value was significant, but it can be justified as an exploratory technique. It is interesting that significant generalization (or differentiation between concepts) was found for the scales *large-small* and *heavy-light* (which have high factor loadings on Osgood and Suci's *potency* dimension), *sweet-sour* (which has a high factor loading on their *evaluative* dimension), and *fast-slow* (which has a high factor loading on their *activity* dimension). Thus, significant generalization occurred along all three of the semantic dimensions proposed by Osgood and Suci. These results have implications for a systematic, quantitative study of semantic generalization.

If three of the four input characteristics used in the learning phase of the present experiment were held constant at some probability level and the other one varied along a set of probability-values, it would be possible to measure systematic changes in semantic generalization as a function of the statistical structure of the concept involved in learning.

(4) *A stereotypic measure.* Another way of obtaining information, although admittedly a limited amount, about a set of conceptual referents would be to ask Ss to describe verbally the most typical member of the set. For the concepts used in the present study, this would involve asking each S to describe the 'stickman' which he felt was typical of the tribe he had just reviewed. This description can be considered to be an index of a stereotype of the tribe. When 37 of the Ss in the probability-learning experiment were asked, after having completed the semantic differential, to describe the typical stickman for the tribe they had just seen, 32 of them gave the most frequently occurring combination of characteristics. This occurred in spite of the fact that this combination appeared only 60% of the time in Tribes A and C and only 40% of the time in Tribes B and D, and it did not adequately reflect the marginal characteristics. The combinations given by four of the remaining Ss were such as to maximize the marginal probability relations, with only one S perceiving some other combination as typical.

SUMMARY

An approach to the study of concept formation and the measurement of meaning was suggested which emphasizes the relationships among (a) the statistical structure of the perceptual characteristics of concept-referents, (b) properties of the mathematical models underlying some devices designed to measure meaning, and (c) the psychological judgmental processes involved in the task of measurement. Experimental data were presented relevant to the process of learning the statistical structure of some concepts and to the measurement of their meanings. The *semantic differential* was used as a measure of meaning, and it was shown that this instrument adequately reflects at least one, but not all, of the statistical aspects of a concept's referents.

PROBABILITY-LEARNING WITH TWO AND THREE CHOICES

By R. ALLEN GARDNER, Army Medical Research Laboratory

Attention has recently been drawn to certain phenomena which have been referred to as "probability-learning."¹ Briefly, the standard experimental situation consists of two or more different stimulus-objects being presented at fixed proportions of the total number of trials, but in random sequences. The subject's (*S*'s) task is to predict which of the objects will appear next on each trial. With random presentation, *S* can only learn the probability of appearance of each object. Thus, he can receive the greatest number of hits by predicting on every trial the stimulus-object which appears most frequently. Nevertheless, in the two-choice situation, *S*s are observed to divide their predictions into proportions which closely match the proportions in which the stimulus-objects are presented. This nets them a proportion of hits considerably below the possible maximum. Yet *S*s have been observed to match their choices to the presented proportions after about 100 trials and to maintain this proportion of responses for as many as 140 subsequent trials.²

It should be noted, however, that the two-choice situation, which has yielded the bulk of this data, is itself an atypical behavioral condition. Usually, behavioral situations both in and out of the laboratory do not restrict *S* to two responses. This is true despite the nearly universal practice of tabulating responses in two classes and treating the data as if only two responses were available. The formulation which follows would seem to bridge the gap which exists between the two-choice and the more typical

* Accepted for prior publication March 10, 1957. Based on a dissertation submitted in June, 1954 to the Graduate School, Northwestern University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author is indebted to Drs. B. J. Underwood and J. W. Cotton.

¹ M. H. Detambel, A test of a model for multiple-choice behavior, *J. exp. Psychol.*, 49, 97-104; W. K. Estes and C. J. Burke, A theory of stimulus variability in learning, *Psychol. Rev.*, 60, 1953, 276-286; W. K. Estes and J. H. Straughan, Analysis of a verbal conditioning situation in terms of a statistical learning theory, *J. exp. Psychol.*, 47, 1954, 225-234; D. A. Grant, H. W. Hake and J. P. Hornseth, Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement, *J. exp. Psychol.*, 42, 1951, 1-5; H. W. Hake and Ray Hyman, Perception of the statistical structure of a random series of binary symbols, *ibid.*, 45, 64-74; E. D. Neimark, Effects of type of nonreinforcement and number of alternative responses in two verbal conditioning situations, *ibid.*, 52, 1956, 209-220.

² Hake and Hyman, *op. cit.*, 64-74.

multiple-choice condition, as well as that between matching behavior and more commonly encountered modes of response which tend to maximize the likelihood of reward.

Suppose two stimuli are presented in the proportions A and B , and S chooses them in the proportions a and b , respectively. Then $A(a)$ will give the proportion of A hits and $B(b)$ the proportion of B hits. Similarly, $A(b)$ and $B(a)$ are the two proportions of misses. Now, when $a = A$ and $b = B$ the two terms representing misses become identical, for they become $A(B)$ and $B(A)$, respectively. This is what takes place when S matches his responses to the presented proportions. It is as if S responds so as to equalize the two proportions of misses (*i.e.* predictions of A when B appears, and predictions of B when A appears).³

Now, in the three-choice situation let A , B , and C be the presented proportions, and a , b , and c the response-proportions. Then hits will be $A(a)$, $B(b)$, and $C(c)$, and misses will be $A(1 - a)$, $B(1 - b)$, and $C(1 - c)$. Setting the error-terms equal to each other again, and solving for a , b , and c we obtain:

$$\begin{aligned} a &= (AB + AC - BC)/(AB + AC + BC); \\ b &= (AB + BC - AC)/(AB + AC + BC); \\ c &= (AC + BC - AB)/(AB + AC + BC). \end{aligned}$$

The response-levels predicted from these equations will not match the presented proportions unless all three are equal. The departure from matching will be in the direction of increased response to the most frequently presented stimulus-object and decreased response to the less frequently presented objects. As a result, the average number of hits will be greater. Furthermore, it can be shown that these effects will increase as the number of choices is increased. Thus, by this mechanism of error-equilibrium the matching behavior found with two choices is seen as a special case, and quite different modes of response are predicted for three or more choices.

These predictions are made by taking the values calculated from the equilibrium-equations to represent average percentages of S 's responses at an asymptote. The values calculated from the equation, however, will be fractions which can vary from -1.0 to $+1.0$, while percentages of response can only vary between $+1.0$ and 0 . Since this will be true whenever responses can only be recorded as positive, it is necessary to determine how predictions can be evaluated when the equilibrium-

³ Employing a somewhat different notation this explanation for the two-choice situation has been pointed out in Grant, Hake, and Hornseth, *op. cit.*, 4.

equation yields negative values for one or more of the choices.

There appear to be two possible positions. First, the equations describe an equilibrium-condition. When S can only make positive responses, and no distribution of positive responses will enable him to equalize his errors, then equilibrium is impossible and we will no longer expect the equations to describe his mode of response. The second position would be to interpret calculated values of zero or less as predictions that the response-level will approach zero or extinction will occur. The operational significance of different negative values will not appear at the asymptote because of this common prediction of zero response. It should be possible, however, to discriminate between different negative levels when further experimental operations are performed such as altering the presented proportions after the responses have reached a zero level. In Experiment I, three-choice combinations were chosen which provided cases where all three of the calculated values are positive as well as cases where one of the calculated values is negative. The second position was chosen as a working hypothesis for this experiment since it yields specific predictions for all cases.

EXPERIMENT I

Apparatus. The apparatus consisted of two or three black metal boxes $4 \times 4 \times 2$ in. placed side by side before S . Set in the top of each box were a small red light and a push-button switch. A large screen placed behind the boxes served to block S 's view of E throughout the experimental session.

Procedure. One light was lit every 4 sec. for 450 trials. The light to be lit was determined by a random series which repeated itself every 150 trials. S was instructed to guess which light would be lit next on each trial. A buzzer indicated the beginning of each trial. S indicated his choice by pressing the button in the appropriate box. If S made 100 consecutive choices of the same light, the experimental session was terminated and the remaining trials scored as if he had continued to choose that light.⁴

Design. The six conditions of the experiment differed in the number and the proportions of lights presented according to the plan shown in Table I. The predictions for final response-levels are also given in Table I. The equations yield negative values for the 10%-choice in the 70-20-10 and 60-30-10 conditions. In both cases predictions of extinction were entered for the 10%-choices for the reasons stated in the previous section. In these cases the predictions for the other response-levels can only be stated as $a > A$ and $b < B$ since the condition $a + b + c = 1$ can no longer be met if c cannot be less than zero.

⁴ This rule was followed in Experiments II and III as well. Its effect was negligible. No S in either of the two-choice groups met the criterion. Only 4.4% of all three-choice S s in all three experiments met it, no more than two cases occurring in any one of the nine three-choice groups.

Subjects. Six groups of 24 undergraduate students served as Ss in this experiment, each group under one condition. The Ss were tested individually and were assigned alternately to conditions in the order in which they arrived at the laboratory. The 70% groups were tested three months before the 60% groups.

TABLE I
DESIGN OF EXPERIMENT I

Presented proportions (%)			Predicted proportions (%)		
A	B	C	a	b	c
70	30	—	70	30	—
70	15	15	80.6	9.7	9.7
70	20	10	>70	<20	0
60	40	—	60	40	—
60	20	20	71.4	14.3	14.3
60	30	10	>60	<30	0

Counterbalancing. Each three-choice group was divided into six equal sub-groups so that all possible spatial arrangements of the boxes were equally represented. For two-choice groups half of the Ss had the A-light on the left and the B-light on the right; this was reversed for the other half.

Results. The results in terms of percentage of response to the various lights are plotted in blocks of 15 trials for the 70-30, 70-20-10, and 70-15-15 groups in Fig. 1, and for the 60-40, 60-20-20, and 60-30-10 groups in Fig. 2. The 70-20-10 and 70-15-15 groups appear to reach a level performance at about Trial 286. Thus, Trials 286 to 450 were taken as asymptotic performance.

TABLE II
MEAN RESPONSE IN PERCENTAGE OF TRIALS 286 TO 450

Group	A	B	C
70-30	72.1	27.9	—
70-15-15	80.2	9.9	9.9
70-20-10	79.8	12.9	7.3
60-40	61.8	38.2	—
60-20-20	67.6	16.2	16.2
60-30-10	68.4	23.5	8.1

The two-choice groups reach final response-levels which very closely match the presented proportions ($P > 0.30$ for both t s). For all three-choice groups the final response-levels are significantly different from the presented proportions ($P < 0.01$ for each t). For the 70-20-10, 70-15-15, and 60-30-10 groups the A-choice response-levels are significantly different from the A-choice response-levels reached by corresponding two-choice groups ($P < 0.05$ for each t). By a one tailed test, $P < 0.05$ for the difference between the 60-20-20 group and the 60-40 group on the A-choice.

The mean response-levels of the 70-15-15 group match those predicted

from the error-equation to within 0.5%, and the response-levels of the 60-20-20 group come to within 3.8% of the predicted levels. Neither difference is significant. The *A*-choice response-levels of the 70-20-10 and 60-30-10 groups are nearly the same as the *A*-choice levels of the 70-15-15 and the 60-20-20 groups.

The final response-level for the 10%-choice was significantly greater than the predicted zero ($P < 0.01$ for both groups) although it was significantly below 10% as noted above. There was no significant difference between the two groups for this choice.

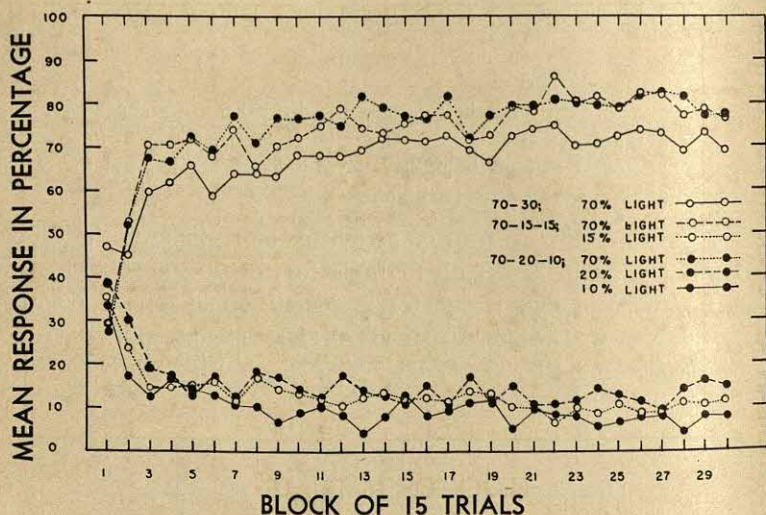


FIG. 1. MEAN RESPONSE IN PERCENTAGE PLOTTED IN BLOCKS OF 15 TRIALS FOR ALL CHOICES OF THE 70-30, 70-15-15, AND 70-20-10 GROUPS

Discussion. The results seem to indicate that the three-choice response-levels are clearly not the matching of presented proportions found with two choices. In the case of the 70-15-15 and 60-20-20 groups the values of the error-equation for all three choices were positive. Here the final response-levels were all fairly close matches of the predicted levels. For the 70-20-10 and the 60-30-10 groups the results are less confirming. In these cases the final levels for the 10%-choices fall far short of the zero levels which would be predicted if negative values of the error-equation are taken to represent extinction. At the same time, the response-levels for the *A*-choices of these groups are very nearly the same as the corresponding response-levels for the 70-15-15 and 60-20-20 groups.

It is possible that the error-equation only predicts for the major choices and that predictions for the low percentage-choices break down unless they are reciprocal as in the case of the 70-15-15 and 60-20-20 groups.

Detambel and Neimark tested groups under similar multiple-choice conditions.⁵ Where their conditions approximated those employed here they found matching behavior similar to that shown by two-choice groups. Detambel, however, terminated his experiment at 200 trials and Neimark at 100 trials, while the 'asymptotes' reported here represent Trials 286 to 450. From Figs. 1 and 2 it can be seen that the results of this experiment

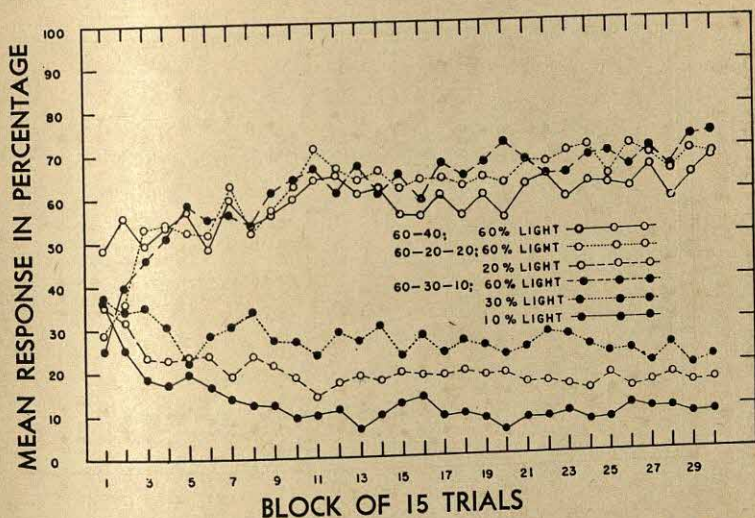


FIG. 2. MEAN RESPONSE IN PERCENTAGE PLOTTED IN BLOCKS OF 15 TRIALS FOR ALL CHOICES OF THE 60-40, 60-20-20, AND 60-30-10 GROUPS

agree with those of Detambel and Neimark up to the point where they terminated trials.

EXPERIMENT II

Problem. In translating the concept of error-equilibrium into algebra as outlined above, an important choice was made in the term used for expressing the error-proportions. With two choices, $A + B = 1$ and $a + b = 1$, there are only two possibilities. Thus, $A = (1 - B)$, $B = (1 - A)$, $a = (1 - b)$, and $b = (1 - a)$. Whether we consider an error as the appearance of one of the lights when it was not predicted— $A(1 - a)$ and $B(1 - b)$ —or the prediction of a light which does not

⁵ Detambel, *op. cit.*, 97-104; Neimark, *op. cit.*, 209-220.

appear— $a(1 - A)$ and $b(1 - B)$ —the results are algebraically equivalent and may be expressed as $A(b)$ and $B(a)$. Behaviorally, both are equivalent also, since S receives the same information whether he predicts a light and is told that he is wrong or he predicts a light and is told that the other one was right.

With more than two choices this is not the case. The set of error-terms $A(1 - a)$, $B(1 - b)$, and $C(1 - c)$ used above is not algebraically equivalent to the other possible set, $a(1 - A)$, $b(1 - B)$, and $c(1 - C)$. Performing the same algebra on this second set of terms yields a new set

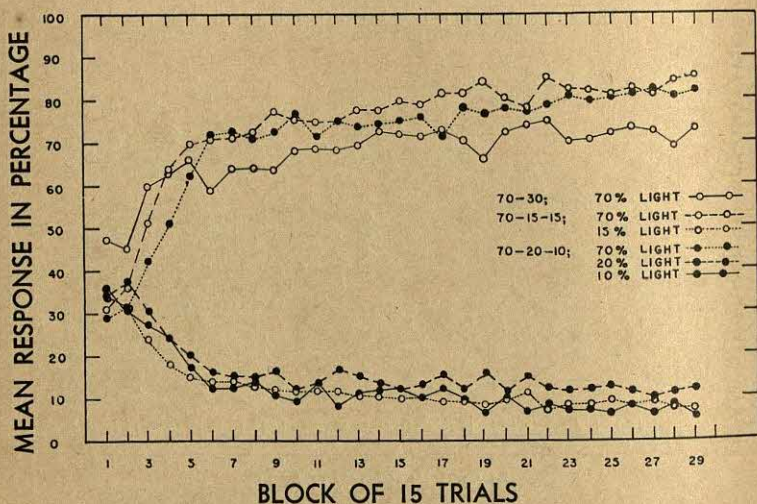


FIG. 3. MEAN RESPONSE IN PERCENTAGE PLOTTED IN BLOCKS OF 15 TRIALS FOR THE CONTINGENT GROUPS OF EXPERIMENT II AND THE 70-30 GROUP OF EXPERIMENT I

of equations. The response-levels predicted from these have properties almost the opposite of those found for the first set. The proportion of responses to the most frequently presented light will be less than the proportions presented; and the proportion of responses to the less frequently presented lights will be greater than the proportions presented. This will yield S a smaller number of hits than matching.

Behaviorally, the difference seems to turn on whether S responds to a stimulus according to the proportion of times that he failed to respond to it when it did occur, $A(1 - a)$; or, whether he stops responding to it according to the proportion of times that it failed to occur after he responded to it, $a(1 - A)$.

Both interpretations appear to apply equally well to the situations described thus far where a stimulus-light appears each time regardless of whether or not it was predicted by *S*. The conditions could be presented, however, in such a way that *S* is only informed of hits and misses and, if he misses, not informed of which light was correct. This latter condition has been referred to as the contingent case, since the information received is contingent upon previous responses.⁶ Similarly, the condition employed in Experiment I is called the non-contingent case. For information as to the 'correct' light is received by *S* on every trial regardless

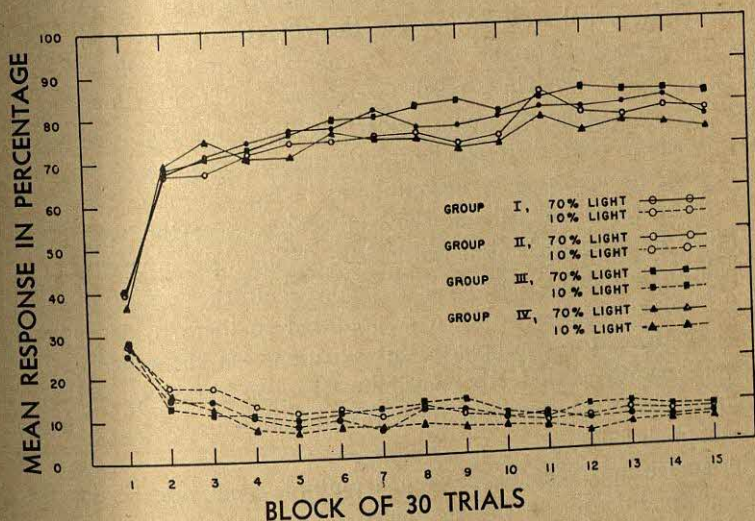


FIG. 4. MEAN RESPONSE IN PERCENTAGE PLOTTED IN BLOCKS OF 30 TRIALS FOR THE 70%- AND 10%-CHOICES OF THE FOUR GROUPS IN EXPERIMENT III

of his previous response. Experiment II was designed to investigate response-levels in the contingent case.

Procedure. The apparatus was identical to that employed in Experiment I. As before, *S* was instructed to guess which would be next on each trial. *S* however, was to expect a light to be lit only after a correct guess. The trials on which a light was lit were determined randomly in such a way that a 70%-light was lit 70% of the time that it was chosen; a 20%-light, 20% of the time; a 15%-light, 15% of the time; and a 10%-light, 10% of the time.

Subjects. Forty-eight *Ss* were used in this experiment; half (24) with the 70-20-10 proportion, and half (24) with the 70-15-15 proportion.

⁶ Detambel, *op. cit.*, 97-104.

Results. An error in recording resulted in the loss of the data for the last 15 trials for many Ss of this experiment. Consequently, complete data were available only for scores up to Trial 435.

The mean percentage of responses to all lights is plotted for both groups in Fig. 3. For our purposes no distinction need be made between the contingent and non-contingent methods of presentation when only two choices are involved. Thus, the 70-30 group of Experiment I was the two-choice group used for comparison with the three-choice groups of Experiment II. As in Experiment I the trials from 286 on were taken as asymptotic performance. The final response-levels of the 70-15-15 group during this period were 82.2% for the 70%-choice and 8.9% for the two 15%-choices. For the 70-20-10 group the final levels were 80%, 12.6%, and 7.4%. These results were not significantly different from those of corresponding groups under the non-contingent conditions of Experiment I. For corresponding intergroup comparisons, the *t*-values were of the same order of magnitude as in Experiment I.

Discussion. As pointed out above, the algebra of the error-equation seems to imply a mechanism which depends upon the appearance of the stimulus-light when it has not been predicted. In the contingent case the light does not appear unless it is predicted so there is some question about applying the equation to this situation. On the other hand, we cannot logically exclude the possibility that the factor at work is S's estimate of the percentage of times that the stimulus is correct when he has not predicted it.

Detambel and Neimark also tested multiple-choice groups under contingent conditions.⁷ As in the non-contingent cases they did not find significant differences between two choices and more than two choices. As in the non-contingent case, however, Detambel terminated at 200 trials and Neimark at 100 trials and at these points the present experiment would have yielded a similar result.

EXPERIMENT III

Problem. In Experiment III two new variables were introduced: distribution of trials and similarity of stimulus-objects. Assuming a mechanism of error-equilibrium, this experiment represents an investigation of the effects of these variables on the operation of the mechanism.

The major deviation of the results of Experiment I from the predictions of the error-equation was the failure of the 10% choice-level of

⁷ Detambel, *op. cit.*, 97-104; Neimark, *op. cit.*, 209-220.

response to reach zero, *i.e.*, response to these lights failed to extinguish as a result of relative lack of reward. Considering the results as incomplete extinction, we might expect degree of extinction to vary with distribution of practice.

In another sense, the error-equation predicts final levels of response which might reflect increased differentiation of response. That is, response-level increases for the major lights and decreases for the minor lights. The stimulus-objects employed in Experiment I consisted of three identical red lights set side by side in three identical black boxes. We might expect that reducing the similarity of the stimulus-objects would tend to increase differentiation in response.

Procedure. The procedure and apparatus were identical to those employed in Experiment I with the exceptions noted below.

Design. Four-sec. trials represented massed practice and 10-sec. trials represented distributed practice. Three identical red lights represented high similarity, and three colored lights (blue, amber, green) represented low similarity. There were four conditions: Group I, 4-sec. trials, identical lights; Group II, 10-sec. trials, identical lights; Group III, 4-sec. trials, different lights; Group IV, 10-sec. trials, different lights. For the low similarity groups the three colored lights appeared in the same spatial order for all Ss, hence color and spatial location were counter-balanced simultaneously.

Subjects. The data of the 70-20-10 group of Experiment I were used for Group I. An additional 72 Ss were used, 24 in each of the three remaining groups also with 70-20-10 proportions.

TABLE III
MEAN RESPONSE IN PERCENTAGE OF TRIALS 286 TO 450

Group	A	B	C
I	79.8	12.9	7.3
II	79.3	12.6	8.1
III	75.5	15.0	9.5
IV	82.7	11.4	5.9

Results. The mean percentage of response to the 70%- and 10%-choices are presented for all fourgroups in Fig. 4. As before Trials 286 to 450 were taken as the final response-level. The final response-levels for all four groups appear in Table III. By analysis of variance final response-levels for the 70%- and 10%-choices yielded significant *F*s for groups with $P < 0.05$. Differences in response to the 20% choice were not significant. Interaction between distribution of practice and similarity was also significant for the 70%- and 10%-choices. The null hypothesis of Bartlett's test for heterogeneity of variance could not be rejected at the 5% level of confidence.

When *t*-tests were performed for differences between the four groups in final response-levels for all three choices, it was found that all differences were insignificant except for two. The differences between Group III and Group IV on both the 70%- and 10%-choices were significant with $P < 0.01$.

Final response-level for the 70%-choice was significantly greater than 70% for all four groups. ($P < 0.01$ in each case). For all except Group III, final level of response to the 70%-choice was significantly greater than the response of the 70-30 group of Experiment I ($P < 0.01$ in each case). Of nine three-choice groups in the three-experiments, only Group III failed to respond to the major choice at a significantly higher level than the corresponding two-choice group. Again, response to the 10%-choice was appreciably greater than zero. It was significantly below 10% ($P < 0.05$) for all except Group III. This was the only case where response to one of the choices was not significantly different from the presented proportion.

Discussion. It is apparent that the major finding of Experiment III was the failure of Group III to reproduce the results of the other 70-20-10 groups and the general results of all of the three-choice groups. Changing from identical lights to different colored lights was sufficient to reduce most of the effects found with all of the other three-choice groups. Yet, when this change was made and trial-length increased simultaneously from 4 sec. to 10 sec., there was no such effect of the colored lights.

Response to the 10%-choice did not extinguish at the values for distribution of practice and similarity of stimulus-objects tested here. This result would seem to force us from the position of interpreting as predictions of extinction negative values calculated from the error-equation. We might retreat to the alternative position that the error-equation will break down when negative responses are necessary for error-equilibrium; or, noting the similarity found in *A*-choice response-levels, it is possible to conclude that the error-equation only predicts accurately for the major choice. In any case, the results for the 10%-choices would seem to preclude a general application of the error-equation to the phenomena of probability-learning in a rigorous fashion.

Whether the error-equation is valid, the results of these three experiments demonstrate that the number of choices influences final response-levels. Any interpretation of probability learning must account for this as well as the matching of stimulus-proportions found under two-choice con-

ditions. This is especially pertinent to the generality which we would wish for such an interpretation. For, it is evident that multiple-choice conditions are closer to those met outside the laboratory.

SUMMARY

A formula was developed to predict levels of response in multiple-choice probability-learning. For two-choice situations it predicts response-levels which match presented stimulus-proportions. For situations involving more than two choices it predicts response-proportions, greater than presented stimulus-proportions for the most frequently presented stimulus, and less than presented proportions for the less frequently presented stimuli.

For two two-choice groups ($N = 24$) response-proportions were not significantly different from presented proportions. For nine three-choice groups ($N = 24$) response-proportions were significantly different from presented-proportions for all three stimuli with the exception of the response to one of the three stimuli for one of the nine groups. For eight of these nine three-choice groups the proportion of response to the most frequently presented stimulus was significantly greater than the proportion of response to a stimulus presented with equal frequency as one of two choices. No significant differences were found between contingent and non-contingent conditions. When similarity of stimuli and distribution of practice were varied a significant interaction effect was found.

THE ROLE OF REPETITION IN ASSOCIATIVE LEARNING

By IRVIN ROCK, New School for Social Research

Although repetition has long been regarded as essential in associative learning, there is some doubt as to how it achieves its beneficial effects. One possibility is that, in learning a list of items, the strength of association between each pair develops gradually, with each repetition adding an increment to the bond, until it is so strong that the first item produces recall of the second. According to this interpretation, repetition is a factor in the *formation* of associations. Another possibility is that repetition is essential because only a limited number of associations can be formed on any one trial. On each new encounter with a list of items we learn new ones, until finally we have learned them all. From this point of view, associations are formed in one trial, and improvement with repetition is only an artifact of work with long lists of items. Typical behavior in experiments on rote learning and in examples from everyday life indicate that some associations are formed in one trial, but such evidence does not prove conclusively that associations are not ordinarily formed by a process of gradual strengthening.

If associations are formed by a process of gradual strengthening based on repetition, it should be easier to form an association between items which have already been presented together on one or more previous trials, but which *S* has not yet been able to get right, than between items presented together for the first time. There are several ways of testing this assumption, but the method used in the experiments to be reported here seems most direct. A control group is given the task of learning a list of paired associates to a criterion of one errorless trial. An experimental group is handicapped by removing all pairs which *S* fails to get right after every trial and substituting new pairs for them. The new pairs are randomly selected from a pool of pairs prepared in advance and from which the initial lists for the two groups also are randomly selected. This means that the experimental group always has the same number of pairs to learn on any given trial—the same number as has the control group—but only some of them will have been seen previously (those already learned) and some will never have been seen previously. Training of the

* Received for publication May 29, 1956. The author is indebted to Mr. Walter Heimer for assistance in this study and to the many students and colleagues who helped in its course.

experimental group also is continued to a criterion of one errorless trial.

For the experimental group, then, a pair is either learned the first time it is seen, or it is removed, and *S* does not, therefore, have what might be presumed to be the benefit of repetition in *forming* associations. If a pair is learned on that first occasion, it will remain in the list as long as *S* continues to get it right. Thus pairs successfully mastered on one trial are repeated, but the repetition does not affect the formation of associations, although *S* gains whatever benefit may accrue from repetition *after* associations are formed. This critical distinction between the role of repetition in forming associations and the role in fixing or reinforcing associations already formed will be discussed later. For the present, it suffices to say that the experimental group has to form associations without benefit of repetition.

EXPERIMENT I

Method. Two groups, of 25 college students each, were required to learn a list of 12 letter-number pairs, which were printed on 3×5 in. cards. The left-hand member of each pair was either a letter of the alphabet or a double-letter. The right-hand member was a number from 1 to 50, since 50 such card-pairs were prepared corresponding to the 25 letters and double-letters, excluding the letter *I*. (The reason for using this type of pair will be discussed below.) The numbers were assigned to the letters by means of a table of random numbers, and the 12 pairs to be used for each *S* were randomly selected by shuffling. In the case of *Ss* in the experimental group, the remaining 38 cards were available as a pool from which new ones could be selected to replace unlearned pairs after each trial. It was thought unlikely that there would be a need for more than 38 substitute-pairs.

A metronome was used to time a 3-sec. exposure of each card and a 5-sec. interval between successive cards. *S* was instructed to associate the letter or letters on each card with the number, and that the serial order of cards would be randomly changed from trial to trial. In view of the procedure used with the experimental group (to eliminate a sense of surprise and lack of understanding of the task which might otherwise be expected), *S* was told in advance that new pairs might be shown from trial to trial, although the total number would remain the same, and that it was his task to try to learn all those shown at any time.

A recall-test for paired associates was used after every trial. On the back of every card was printed only the first item of the letter-number pair. The 12 cards used in that learning trial were therefore reversed, shuffled, and presented at the rate of one every 5 sec., *S*'s task being to respond with the correct number. In the case of the experimental group, the cards were sorted in two piles as *S* responded, depending upon whether he was right or wrong. New cards were then substituted for the wrong ones, and these were shuffled together with the right ones in preparation for the next trial. There was a 30-sec. interval between each test-trial and the subsequent learning trial.

After about three-fourths of the *Ss* had been tested, it was noted that the average first-trial performance of the experimental group was higher than that of the control group. To correct for this sampling error, the remaining *Ss* were not assigned

to a group until they had completed the first trial. Those doing relatively well (four or more pairs correct) were then assigned to the control group and those doing relatively poorly (three or fewer correct) were assigned to the experimental group. Since the procedure is the same for both groups for the first trial, this was a perfectly good way of equating the groups. The result was that the two groups as finally constituted had virtually identical scores on the first trial.

Results. It turned out that the number of extra pairs available for substitution in the lists of the experimental group was not sufficient in every case. For five Ss, the experiment had to be discontinued because there were insufficient new cards to substitute for unlearned pairs on the last trial given. These Ss were scored 10+ which means that they would have required at least 10 trials to reach the criterion. Similarly, there were three Ss in the control group who did so badly that work with them was discontinued before the criterion was reached, and they too were scored 10+. The median number of trials to reach the criterion for the two groups was exactly the same; namely, 4.75. The semi-interquartile range also was identical for the two groups; namely, 1.45. If the cases scored as 10+ are excluded, the mean of the control group is 4.55 and the *SD* is 1.9, while the corresponding values for the experimental group are 4.35 and 1.2. As can be seen by inspection, the mean difference falls far short of statistical significance ($t = 0.41$). The mean number of errors (no response *plus* wrong answers) to criterion was 17.9 for the control group ($SD = 9.9$) and 17.2 for the experimental group ($SD = 8.6$), and this difference too lacked statistical significance ($t = 0.24$). A plot of the mean number of correct responses per trial yields highly similar curves for the two groups. There was thus no advantage for the control group.

The reason for using letter-number pairs instead of other material, *e.g.* nonsense-syllable pairs, in the first experiment was as follows. It was thought that, if nonsense-material were used, the control group might have an advantage unrelated to the hypothesis under investigation. Repetition for this group would certainly have had the effect of making each individual item familiar, and there was reason to suppose that it would be easier to form associations between familiar items than between unfamiliar ones, even though the familiar items had never in the past been associated with each other. Since one does not have to learn the items as such if they are familiar, one can concentrate on the associations to be formed. If nothing else, it ought to be easier to *recall* a familiar item than an unfamiliar one. By using pairs of letters and numbers which were already familiar, an attempt was made to eliminate this factor. Of course, even here one might argue that the control group had an advantage because of

the additional familiarization in successive trials with the material used. Nevertheless, the use of familiar material would minimize this factor, and the results suggest that, even if there were such an advantage, its effect was negligible. In using familiar items, however, the problem arises of finding material in sufficient quantity and in other respects suitable for this type of experiment. By using letters and double-letters with numbers, 50 pairs were obtained but, as noted above, even this number turned out to be insufficient. Primarily to overcome this shortcoming, and in spite of the objection mentioned above, it was decided to repeat the experiment with pairs of nonsense-syllables. There is also something to be said for using this more traditional material because of the greater possibility afforded for comparison with previous work.

EXPERIMENT II

Method. The procedure followed that of Experiment I, except that pairs of three-letter nonsense syllables were used. They were taken from Glaze's list of syllables of 47% and 53% association value.¹ Eighty pairs were made up, from which 8 were randomly selected as the starting list for Ss of both groups, and there remained a pool of 72 pairs from which pairs could be randomly drawn as needed to substitute for unlearned pairs in the case of the experimental group. A response was scored correct if S pronounced the syllable correctly, and partially correct responses were scored as wrong. It was not necessary in this experiment to assign Ss to one or the other group after determining their scores for the first trial, because it turned out that performance on this trial was not significantly different for the groups as they were randomly constituted. There were 15 adults of both sexes, mostly college students, in each group.

Results. In addition to the 30 Ss, there were 8 Ss who did not succeed in learning the list to the criterion, 3 in the control group and 5 in the experimental group. In all these cases, E discontinued the experiment because it was obvious that there was little likelihood that S would master the task in the time available. Several of these Ss were evening-session students tested at night, after work, when they were tired. In a few cases, the students had to return to a class, and therefore could not continue to the criterion, and two were somewhat older people. The pool of 72 pairs turned out to be more than adequate and had nothing to do with the discontinuation of the experiment for the Ss of the experimental group.

Considering first the results for the 15 Ss of each group who did learn, the mean number of trials to reach the criterion turned out to be identical,

¹ J. A. Glaze, The associative value of nonsense syllables, *J. genet. Psychol.*, 35, 1928, 255-267. Cf. also E. H. Hilgard, Methods and procedures in the study of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 541-544.

namely, 8.1. The *SD* was 3.0 for the control group and 2.4 for the experimental group. The mean number of errors to criterion was 26.7 for the control group (*SD* = 11.4) and 29.2 for the experimental group (*SD* = 12.6). The difference is not statistically significant ($t = 0.55$). If the eight *Ss* who failed to reach the criterion are included, scoring them as 15+, the median number of trials also is identical for the two groups; namely, 8.5. The semi-interquartile range is 3.2 for the control group and 4.2 for the experimental group. The learning curves for the two groups are again similar.

DISCUSSION

There were reasons for expecting superior performance in the control groups of these experiments, entirely apart from the question of repetition. The possible advantage of familiarity, which applies particularly to Experiment II, has been mentioned. A second reason is that a recall test rather than a matching test of recognition was used. In a matching test, *S* must recognize which item belongs with which, but he does not have the additional task of recalling the second item. He need only recognize it. It is, therefore, an easier task, and it is reasonable to suppose that the strength of an association must be greater for recall than for matching. If so, there must have been many instances in which the *Ss* of both groups were scored wrong but would have succeeded in a matching test. In fact, the many cases where *S* gave a partially correct response (almost the right number in Experiment I or two of the three letters correct in Experiment II), but was scored wrong, point in this direction. No doubt in such cases as well, *S* would have been correct if he had the advantage of a matching test. This would mean that an association did in fact exist, and the *Ss* of the experimental groups were penalized by having such pairs eliminated. A matching test was not used because it entailed certain technical difficulties in the case of the experimental groups in view of the unpredictable change of items from trial to trial. That the control groups were not superior despite this possible advantage, however, makes the case against the incremental theory all the stronger.

Another reason that one might have expected superior performance of the control groups relates to the fact that wrong pairs were removed from the list for the experimental groups even if they had been correct on one or more previous trials, although that happened only occasionally (primarily during the first few trials). Since it may be assumed that an association has been formed once *S* gets a pair correct, it was not necessary to remove these pairs to test the hypothesis under consideration. Doing so conse-

quently constituted something of a very strict procedure for the experimental groups. When the same thing happened to an *S* in the control group, he had what is no doubt the advantage of being presented with the same pair again. If the association continued to exist—and failure on the previous trial was a matter of momentary forgetting—he would not really have to learn the pair. The *S* in the experimental group, however, had to learn a new pair to take its place. Why the control groups do not learn more rapidly in view of these reasons is not clear, except that apparently they do not add up to an advantage of strength sufficient to reveal itself under the conditions employed.

There is one other difference between the conditions of the two groups which, however, might be said to favor the experimental group. It may be argued that for every *S* there are certain pairs which are easier to learn than others. That being the case, the easy pairs may be the ones which are learned on any given trial, while the difficult ones are eliminated by a process of 'natural selection.' The new pairs substituted are not necessarily easy, but the reconstituted list on the whole is not as difficult as that with which the control group is faced on a corresponding trial. It is not easy to deal with this objection experimentally if difficulty is defined idiosyncratically, because the only way of finding out about a pair is to present it to that *S* for learning. Work on this problem is now in progress. Even if it should turn out, however, that the difficulty factor does work to the advantage of the experimental group, the meaning of such an advantage must be seen in the proper perspective. As things now stand, there are the several factors, mentioned above, which tend to favor the control group, and which ought to counterbalance any such advantage, but even if control of difficulty were to impair the performance of the experimental group to some extent, the question of interpreting the magnitude of the superiority of the control group would remain. One might still be struck by the level of performance achieved by the experimental group without benefit of repetition.

The present results seem to support the thesis that, in the classical multiple-item learning situation, associations are formed in one trial. Informal observations and introspective reports support this finding. Most *Ss* succeed in learning only a few pairs per trial, and many do so with the aid of some mnemonic device. The theoretical significance of the widespread use of such devices in rote learning experiments has not been sufficiently emphasized in the past. The successful use of such devices may mean that an idea suddenly occurs to *S* which enables him to link two items then and there; it has, to some extent, the character of insightful

learning. Some Ss concentrate on one or two pairs and hardly even attend to the rest; other Ss do try to learn them all. With or without the use of devices, the difficulty in getting more than a few pairs correct in any one trial has to do with the presence of so many other pairs.

It follows from the argument that repetition is important only for the learning of *additional* associations that in situations with few associations to be formed, repetition would not be required at all. This deduction can easily be tested by presenting Ss with only one pair of items (under conditions of incidental learning which eliminate the possibility of 'silent' repetition).²

Why, in multiple-item learning situations, it is not possible as a rule to learn more than a few associations on any one trial, remains to be investigated, but the following point must be considered: If, a few seconds after any given pair is exposed during learning, and before the next pair is shown, *E* were to test for retention, there is little doubt that most Ss would give the correct response. Would this result not mean that an association was in fact formed, and that the failure to get all items right on a test following the entire series is a matter of forgetting of already formed associations? In that event, one could regard forgetting due to the presence of many items as an instance of retroactive inhibition. According to this conception, repetition would not be required to *form* associations, even in the case of long lists of items, but only to strengthen them sufficiently to resist interference. The present experiments rule out this type of explanation, since Ss do as well with substituted new pairs as with the original ones. It is as if pairs which are not retained by the time of the test leave nothing in the nervous system of any value for future use. The Ss frequently comment spontaneously that they do not recognize the items of a pair which they have seen several times before.

The question is really one of deciding on a criterion for the existence of a formed association. It probably makes more sense—in terms of what we usually mean by associative learning—not to define success by the easy test of immediate memory, but rather by the traditional delayed test. At any rate, it may be said that repetition is required in rote learning, as defined by this latter criterion (the traditional one), because for some as yet unknown reason or reasons the existence of many items interferes with the formation of more than a few associative connections of a nature which is of any value beyond immediate memory. The hypothesis of Von Restorff

² Preliminary results reported by F. B. Springfield (The learning and retention of a single pair of associates, B.A. thesis, New School for Social Research, 1955) indicate that almost all Ss do succeed in establishing an association in one trial.

concerning loss of individuality of traces under conditions of crowding is relevant here, but leaves many questions unanswered.³

It seems clear that repetition plays another role in rote learning; namely, that of strengthening or reinforcing associations once they are formed. There is little question that, in everyday life, we have strong associations between items which have been experienced contiguously very often. Experiments on overlearning clearly demonstrate this point, but they do not distinguish between the role of repetition before associations are formed and after they are formed, the impression being that repetition serves continuously to build up the strength of the association from zero-value to a value far in excess of that required for initial recall on the test following the learning trial. The present results show that this is not the case; repetition does not seem to be of value in forming associations. Hence, it must be concluded that, in overlearning, only the repetition *after* the association is formed is effective in strengthening it.

The present findings are certainly compatible with the thesis that an association is essentially the after-effect of an organization of the items during the initial experience.⁴ Interestingly enough, however, the exponents of this point of view have never denied the importance of repetition in associative learning, tending to believe that, in the case of nonsense material in monotonous series (as contrasted with meaningful material), repetition is needed to consolidate what would otherwise be very unstable or weak trace-interrelationships.⁵ Thus, in the case of rote learning it was implicitly taken for granted that associations are formed by a process of gradual strengthening. On the other hand, the present findings are incompatible with incremental theories, including those which stress the law of effect.

SUMMARY

In two experiments, the learning of paired associates was studied. In each case, the traditional procedure was used for a control group, while for the experimental group unlearned pairs were removed and new ones submitted after each trial. No significant differences in rate of learning were found. This result suggests that repetition plays no role in the formation (as distinct from the strengthening) of associations, other than that of providing the occasion for new ones to be formed, each on a single trial.

³ Hedwig von Restorff, Ueber die Wirkung von Bereichsbildung im Spurenfeld, *Psychol. Forsch.*, 18, 1933, 299-342.

⁴ Wolfgang Köhler, On the nature of associations, *Proc. Amer. Phil. Soc.*, 84, 1941, 492.

⁵ Kurt Koffka, *Principles of Gestalt Psychology*, 1935, 556-571.

ALTERNATIVE MEASURES FOR THE DISCRIMINATION OF SHIFT IN REINFORCEMENT-RATIO

By ALLEN PARDUCCI, Swarthmore College

Changes in performance following shifts in schedules of reinforcement provide significant but difficult targets for general theories of learning. The paradoxically greater resistance to extinction following partial as compared with 100% reinforcement has been 'explained' in both cognitive and S-R terms. The alternative approaches share an emphasis upon the discriminability of the shifts, but they differ in their interpretations of what is being discriminated.

Humphreys hypothesized that *Os* discriminate the *regularity* of the reinforcement schedules.¹ The similarity between the pre- and post-shift schedules on this dimension of regularity was assumed to determine *Os*' expectancies and consequently their resistance to extinction. Sheffield presented evidence in support of an analysis which based resistance to extinction upon the similarity of the immediate stimulus-elements for each of the pre- and post-shift trials.² She argued that discrimination could be more univocally deduced from the principle of *stimulus-generalization* (and a piecemeal analysis of the stimulus-situation) than from the *ad hoc* relationships between regularity and expectancy postulated by Humphreys. Bitterman and his associates have demonstrated the inadequacy of Sheffield's trial-by-trial analysis and have proposed instead that the crucial discrimination may, under certain circumstances, be made between the *pattern* or *sequence of events* before and after the shift.³ Grant also has emphasized the overall schedule or rate of reinforcement.⁴ He has suggested, furthermore, that in human betting experiments discrimination of this rate may be reflected in *Os*' overall rate of response even though that rate may not appear to be directed toward maximization of reinforcement.⁵

* Received for publication July 31, 1956.

¹ L. G. Humphreys, Acquisition and extinction of verbal expectations in a situation analogous to conditioning, *J. exp. Psychol.*, 25, 1939, 294-301.

² V. F. Sheffield, Resistance to extinction as a function of partial reinforcement and distribution of practice, *J. exp. Psychol.*, 39, 1949, 511-526.

³ Janet Crum, W. L. Brown, and M. E. Bitterman, The effect of partial and delayed reinforcement on resistance to extinction, this JOURNAL, 64, 1951, 228-237; E. D. Longenecker, John Krauskopf, and M. E. Bitterman, Extinction following alternating and random partial reinforcement, *ibid.*, 65, 1952, 580-587; D. W. Tyler, E. C. Wortz, and M. E. Bitterman, The effect of random and alternating partial reinforcement on resistance to extinction in the rat, *ibid.*, 66, 1953, 57-65.

⁴ D. A. Grant, H. W. Hake, and J. P. Hornseth, Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement, *J. exp. Psychol.*, 42, 1951, 1-5.

⁵ Grant, Information theory and the discrimination of sequences in stimulus events, in *Current Trends in Information Theory*, 1953, 18-45.

The nature of the preshift-postshift discrimination is inferred, in the above approaches, from the resistance to change exhibited by the respective *O*s following different kinds of stimulus-shift. With human *O*s, independent indices of the nature of the discrimination may be obtained by asking for different kinds of discrimination and comparing the answers with *O*s' resistance to change. Grant has shown that questions about past reinforcement rates can provide evidence for a precision of discrimination not shown by conditioned responses.⁶ The present experiments use this approach to explore the nature of the discriminations occurring in a Humphreys-type, verbal-conditioning situation. They take off from an experiment by Estes and Straughan in which differences in adjustment to a postshift schedule of reinforcement appeared to be associated with the difficulty of discriminating between the pre- and post-shift schedules.⁷

EXPERIMENT I

The first aim of Experiment I was to replicate the essentials of Estes and Straughan's two-choice betting experiment. The phenomenon of special interest in this context—the more complete postshift adjustment by the group shifted from a 15-85 to a 70-30 ratio than by the group shifted from 50-50 to 70-30—was of only borderline significance in the original experiment. The second aim was to determine in what way differences in postshift adjustment were associated with differences in *O*s' discrimination of the postshift schedule. Half the *O*s in each experimental group were asked to make periodic estimates of the frequencies with which the alternative stimulus-objects had appeared *in the past*. The remaining *O*s were asked to estimate the frequencies with which the stimulus-objects would come *in the future*. The hypothesis was that there would be significant differences between the past and future estimates made by the respective groups, and that the postshift betting rates would be more closely associated with differences in future than in past estimates.

Procedure. The method used by Estes and Straughan was modified for purposes of economy. Instead of flashing lights, the stimulus-objects were the letters *A* or *B* read aloud to the respective groups. All *O*s were first given response-booklets and were told that they were being given a test to measure how well they could learn the relative frequencies with which different letters appear. Their specific task was to anticipate, for each successive trial, which letter would be read next and to record it in the appropriate space in their booklets.

Instructions. After the 80th bet, *O*s were instructed to turn over the page and to follow the instructions printed on the next page. Half of each group, the estimators of past frequency, read the following:

You have thus far been given 80 different presentations of the letters. Without

⁶ *Ibid.*, 37-39.

⁷ W. K. Estes and J. H. Straughan, Analysis of a verbal conditioning situation in terms of statistical learning theory, *J. exp. Psychol.*, 47, 1954, 225-234.

turning back to the first page, try to make as accurate an estimate as you can of how many of these 80 were of *A* and how many were of *B*. Record your estimates in the following blanks.

For the other half, the estimators of future frequencies, the instructions read:

You have thus far been given 80 different presentations of the letters. By this time you probably have acquired some expectations about the relative frequencies with which *A* and *B* will be given during the next 80 presentations. Without turning back to the first page, try to predict as shrewdly as you can how many of the next 80 will be *A* and how many will be *B*. Record your estimates in the following blanks.

Similar instructions appeared after the 160th and after the 240th, or last, presentations (an extra page for bets being attached at the end that the request for future estimates would not seem inappropriate). Each *O* received the same kind of instructions on all three frequency-estimation pages.

In other ways, the experimental conditions approximated those of Estes and Straughan. The same preshift frequencies were used; in this case, *A* was correct on 15% of the 120 preshift trials for one group (40 *O*s), on 50% of these trials for the second group (38 *O*s), and on 70% for the control group (25 *O*s); for all three groups, *A* was correct on 70% of the 120 postshift trials. There was no indication of a break at the point of shift.

Results. The actual betting frequencies are graphed in Fig. 1. As is the irrational custom in such experiments, *O*s' bets approximated the actual stimulus-frequencies (rather than the maximal gain which would involve either 100% or 0% bets on *A* depending upon whether *A* was discriminated as appearing for more or less than half the trials). Irregularities in the curves may have been enhanced by the fact that the same randomized series was presented to all members of each group and thus may reflect unusual runs (the experimental *A/B* proportions being maintained only for each successive set of 40 trials).

The direction of the difference between the mean bets of the 15% and the 50% groups during the postshift series (for which the stimulus-order was the same for all groups) is the same as was found by Estes and Straughan. An analysis of variance performed upon the sum of the *A*-bets during the last 80 postshift trials (following the procedure described for unequal *N*s by Walker and Lev)⁸ yielded a significant mean square for differences associated with preshift reinforcement ratios ($F = 9.21$; $df. = 2$ and 97 ; $p < 0.001$). Differences associated with the nature of the interpolated estimates (past or future) and with the interaction between these and preshift ratios were practically nonexistent. Thus there is real evidence that adjustment to the post-shift schedule of reinforcement is more com-

⁸ H. M. Walker and Joseph Lev, *Statistical Inference*, 1953, 381-382.

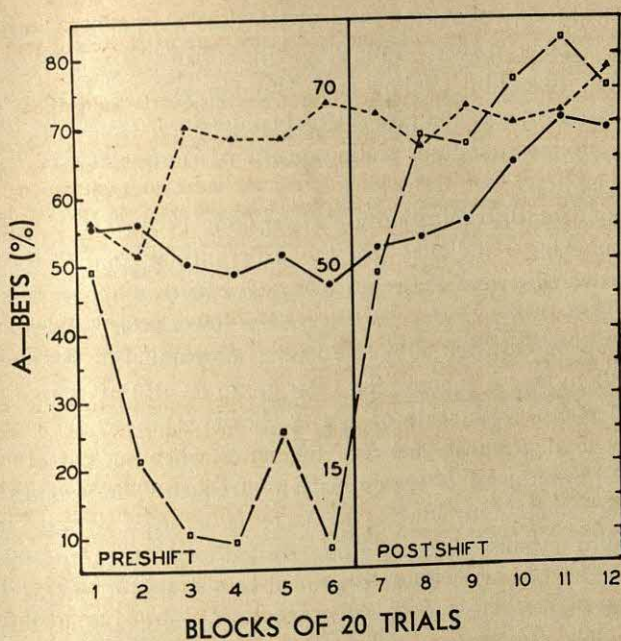


FIG. 1. MEAN BETS ON A IN EXPERIMENT I

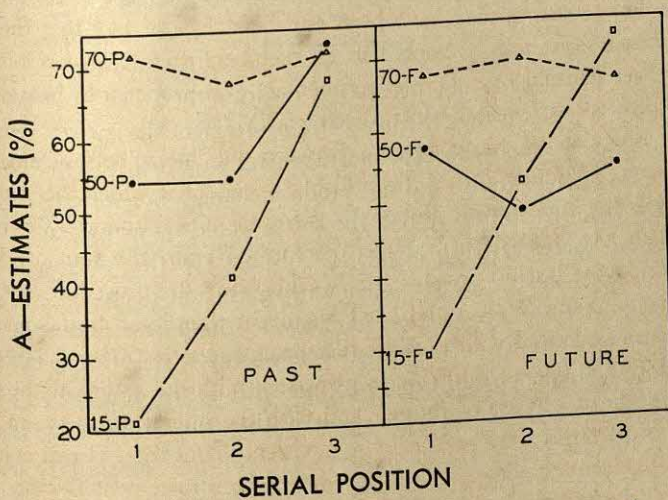


FIG. 2. MEAN ESTIMATES OF PAST AND FUTURE FREQUENCIES OF A IN EXPERIMENT I

plete when the shift from the original schedule has been greater. This is what one might expect following Bitterman's *similarity-of-pattern* analysis; but it would not follow from Humphreys' *similarity-of-regularity* analysis since both pre- and post-shift schedules are irregular for all groups in this arrangement, and any scaling on the basis of regularity would appear to be arbitrary.

What kind of discrimination are *O*s making between the series? The left side of Fig. 2 indicates that the experimental groups discriminate quite effectively between the series *when such discrimination is measured in terms of the past-frequency estimates*. The lower post-shift betting level by the group shifted from 50% to 70% has no parallel in this measure of discrimination; here it is above the 15% group on all three past estimates. As shown on the right side of Fig. 2, however, the 50% group fails completely to discriminate between the series when such discrimination is measured in terms of *future* estimates of frequency, and its final future estimates are significantly lower ($t = 3.21$; $df. = 36$; $p < 0.01$) than the corresponding estimates by the group shifted from 15% to 70%. (An overall analysis of variance was precluded by the almost total absence of variance between the past estimates of the 50% group.) Thus the slower postshift adjustment following the less dramatic stimulus-shift appears to be associated, not with a failure to discriminate the occurrence of the shift, but rather with a failure to discriminate its *permanence*.

EXPERIMENT II

The fact that *O*s do not necessarily readjust their betting when they discriminate a shift in the schedule of reinforcement fits in well with the results of experiments with absolute judgment in which it appeared that shifts in the stimulus-range did not produce changes in scale of judgment unless *O*s were led to believe that the shifts would be permanent.⁹ It was hypothesized there that such beliefs would be a function of how likely a sample the post-shift series was from an hypothetical population with the characteristics of the pre-shift series. Thus in the first experiment, since the 70-30 post-shift series was a more likely sample, statistically speaking, from the 50-50 than from the 15-85 pre-shift series, the group shifted from 15-85 was more ready to believe that the shift was not just temporary.

⁹ Allen Parducci, Direction of shift in the judgment of single stimuli, *J. exp. Psychol.*, 51, 1956, 169-178; Allen Parducci and Raymond Hohle, Restriction of range in the judgment of single stimuli, this JOURNAL, 70, 1957, 272-275.

Procedure. Experiment II was designed to provide a test of this interpretation and also to further explore the effects upon discrimination of different degrees of shift in the schedule of reinforcement. Again, all groups had a postshift level in which *A* occurred for 70% of the trials. One group was shifted from a 40% and a second from a 60% level. In accordance with the above interpretation, the former group should adjust more readily to the post-shift level. A third group was shifted from 30-70 to 70-30 in order to test an alternative hypothesis more consistent with the *learning-set* or strategic approach favored by Goodnow and Pettigrew.¹⁰ This hypothesis interprets the results of Experiment I as following from the fact that the 15% *Os* learn in the pre-shift series that *A* and *B* come with *unequal* frequency, and thus they find it easier to learn in the post-shift series that they will come with unequal frequency, but with the ratio reversed. (Note that the proportion of *A*-bets was actually higher than the control group's during the last 80 trials.) Since

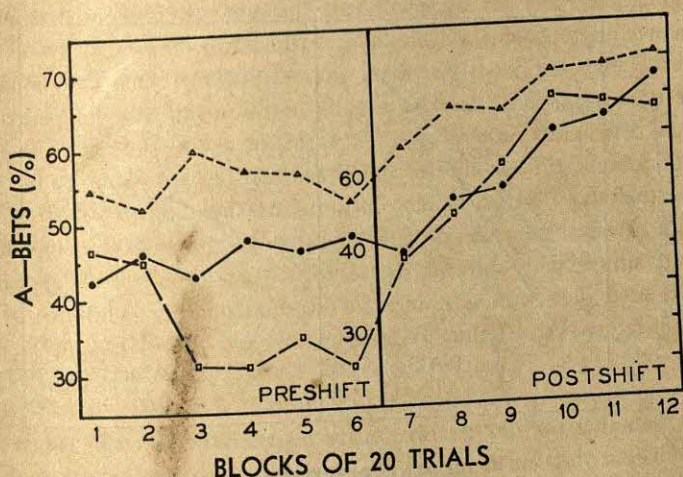


FIG. 3. MEAN BETS ON *A* IN EXPERIMENT II

the ratio for this new group is exactly reversed, it should adjust to the post-shift level more readily than any of the other groups—or at least as quickly as the 15% group in Experiment I. Once again, independent measures of *Os*' discriminations were obtained using both the future and past estimates of frequency. The 30%, 40%, and 60% groups were composed of 20, 34, and 25 new *Os*, respectively. As in Experiment I, all *Os* were college sophomores.

Results. The differences in post-shift adjustment, plotted in Fig. 3, are less dramatic than in Experiment I, but, the analysis of variance performed in the same way for these new groups indicated significant differ-

¹⁰ J. J. Goodnow and T. F. Pettigrew, Effect of prior patterns of experience upon strategies and learning sets, *J. exp. Psychol.*, 49, 1955, 381-389; Goodnow and Pettigrew, Some sources of difficulty in solving simple problems, *ibid.*, 51, 1956, 385-391.

ences associated with preshift ratios ($F = 3.75$; $df. = 2$ and 73 ; $p < 0.05$). Again, neither the effects of the interpolated estimates of frequency nor the interaction between these and the pre-shift ratios were significant. The interesting fact, however, is that it is the 40% rather than the 60% group which shows most resistance to change; its level of readjustment (as measured by the number of *A*-bets made during the last 80 trials) is lower even than that of the 50% group of Experiment I. These findings are not consistent with the sampling hypothesis that resistance to shift increases with the similarity between the pre- and post-shift schedules.

Nor does the slow readjustment by the 30% group provide any support for the assumption that these *O*s should learn to expect the alternatives

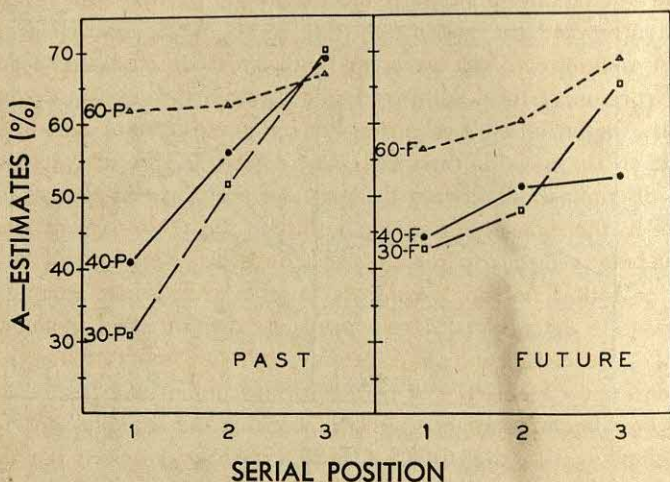


FIG. 4. MEAN ESTIMATES OF PAST AND FUTURE FREQUENCIES OF *A* IN EXPERIMENT II

to come with unequal frequency and should thus adjust to the post-shift level more rapidly. If anything, the post-shift level of the 30% group is even lower than that of the 50% group and much lower than that of the 15% group. An overall analysis of variance performed on the post-shift scores of all 12 subgroups indicated that the differences associated with the pre-shift ratios were significant at the 1/10% level.

How do the new groups differ in regard to the other measures of their discriminations between the series? The left half of Fig. 4 indicates that there is again no difficulty in discriminating that a shift has occurred, but the right half suggests that at least one of the groups had doubts about the permanence of the shift. An analysis of variance performed on the

third future estimates of frequency of all six groups was significant ($F = 3.82$, $df. = 5$ and 84 ; $p < 0.01$), and the application of Tukey's gap-test at the 5% level indicated that a significant gap separated the 40% and 50% groups from the remaining four. While a corresponding analysis for the past estimates was precluded by the absence of variance in the 50% group, the t between the lowest mean (60% group) and the next to the highest (70% group) was less than 1.0, indicating that none of the differences in final past estimates were significant.

DISCUSSION

When the results of both experiments are combined, we find a curvilinear relationship between the magnitude of the shift in ratio of reinforcement and the postshift betting levels of the respective groups. Greatest resistance to change occurs with shifts of moderate degree. A similar curvilinear relationship is found between the degree of stimulus-shift and the final *future* estimates, but the groups cannot even be differentiated on the basis of their final *past* estimates. This suggests that the crucial discrimination between the pre- and post-shift series does not depend upon the difficulty of noting a shift in the reinforcement schedule. The differences in the expectancies concerning the *permanence* of the shifts appear to follow no simple rule. As in other areas where attempts have been made to describe the development of expectancies in common-sense terms (e.g. irrelevant-incentive learning, continuity-noncontinuity), common sense must be elaborated before it can handle the experimental evidence. *Post hoc*, one might speculate here that the stimulus-shift for the 60% group was so slight that its final betting level reflects not the discrimination of a permanent change in the schedule of reinforcement but rather a subtle sharpening of what they believe to be the very same schedule.

Within an S-R framework, the results cannot be handled using the principle of stimulus-generalization alone (even by defining the stimuli in terms of the overall pattern of reinforcement) since gradients of generalization have always been assumed to be monotonic—an analogy with rote learning might have prevented such optimism. Applying Osgood's resolution of the similarity paradox in transfer and retroaction to the present data,¹¹ one could assert that the smaller the numerical difference between a group's pre- and post-shift schedules of reinforcement, the

¹¹ C. E. Osgood, The similarity paradox in human learning: A resolution, *Psychol. Rev.*, 56, 1949, 132-154.

greater the similarity between its pre- and post-shift *stimuli*; but since stimuli and responses are identical here (assuming an Estes-type analysis in which the crucial response is not *O*'s bet but rather his unconditioned reaction to the reinforcing stimulus),¹² increased stimulus-similarity is paralleled by increased response-similarity. Plotting the hypothetical changes for these two dimensions on Osgood's tridimensional transfer surface, one would move diagonally from the corner representing stimulus-dissimilarity and response-antagonism toward the corner representing stimulus- and response-identity. The change in transfer would be first down and then up, corresponding to the results of the present experiments.

Neither the cognitive nor the *S-R* approaches specify the point of reversal. It is interesting that while the 40% group shows most resistance to postshift adjustment in terms of both the betting and estimate of future frequency, the 30% group is lower than the 50% group in its betting but higher in its estimates. It should be mentioned also that there is little correlation between these measures for *O*s within each group or even between their second and third future estimates. This suggests that their discriminations or expectancies are of a transitory nature, varying considerably with the serial position as well as with the manner of their measurement. The group-differences indicate, however, that extreme resistance to readjustment here is not due to failure to discriminate *any* shift in the schedule of reinforcement, but rather to a failure to grasp the permanence of such shift. The conditions determining this latter kind of discrimination remain to be elaborated.

SUMMARY

Betting behavior and periodic estimates of both past and future frequencies of stimulus were studied in a Humphreys' two-choice guessing game. Alternative randomized series, with reinforcement ratios of 15-85, 30-70, 40-60, 50-50, 60-40, and 70-30, respectively, were shifted to a common, 70-30 series. Betting adjusted to the actual postshift ratio most completely following either the greatest or the smallest stimulus-shifts. This curvilinear relationship between the changes in reinforcement- and betting-ratios was associated with estimates of future stimulus-frequencies but not with estimates of past frequencies. The crucial discrimination between the series thus appeared to be concerned with the permanence of the shift and not with the occurrence of the shift *per se*.

¹² Estes and Straughan, *op. cit.*, 225-234.

THE TARTINI PITCHES CREATED BY THE PRIMARY NUMBER SEVEN

By MAX F. MEYER, Miami, Florida

Neither Tartini nor Helmholtz, investigating 'subjective tones,' experimented with intervals outside of the triple-triad diatonic scale. The only prime numbers as ratio-terms used for their sensory experiments by Tartini and Helmholtz were the prime numbers 1, 2, 3, and 5. The consequence for the theory of hearing has been serious—assumptions about difference and summation tones have been made and perpetuated that are not justified.

In a very ancient Greek epoch musical theorists admitted only the prime numbers 1, 2, and 3, which we now call the Pythagorean system. In later Greek times the four prime numbers 1, 2, 3, and 5 were used as factors. At the end of the Middle Ages the theorist Zarlino (1517-1590) thought he had a special limiting reason for that; viz. "Music must not get outside the *six dimensions* of space: fore, aft, right, left, up, down." The Chinese, although not influencing the western world, also had a reason, because for them the prime number 5 was a sacred number, viz. north, east, south, west, and center. Under such influences Tartini (1692-1770), in his experiments with subjective 'notes' made by violins and oboes, failed to use the prime number 7 as a ratio-term. Helmholtz did the same without explaining why.

The great mathematician Leonhard Euler (1707-1783) had no respect for such reasons and freely used the numbers 1, 2, 3, 5, and 7 in his musical theory. We should do the same in the physiological theory of hearing. Indeed, for a purely physiological theory of hearing we must go beyond all restrictions which an esthetic theory of music imposes. We might, however, begin with the tritones, which by the medieval theorists of music were regarded as horrible, almost devilish.

Positions of the tritones within the octave span, between two well-known intervals, are as follows:

- 3:4, Pure fourth, equals 4980 mills of $1/12 \log 2$
- 5:7, Minor tritone, equals 5825 mills of $1/12 \log 2$
- 7:10, Major tritone, equals 6175 mills of $1/12 \log 2$
- 2:3, Pure fifth, equals 7020 mills of $1/12 \log 2$

Major tritone. To study the Tartini pitches of the tritones, I proceed as follows.

* Received for publication November 11, 1955.

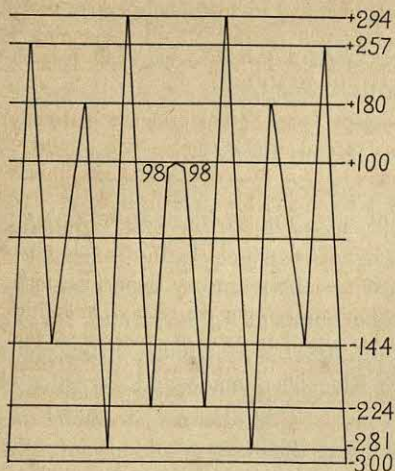


FIG. 1. HYDRAULIC EFFECT OF
 $2\cos 7x + \cos 10x$

Only 7 of the possible maxima appear; Pitch 10 is masked. $224 + 98 = 322$ and $144 + 180 = 324$. One-half of these virtually identical hither-thither motions is 162. And ± 162 absorbed (shaved off) from every maximum and minimum transforms Fig. 1 into Fig. 2.

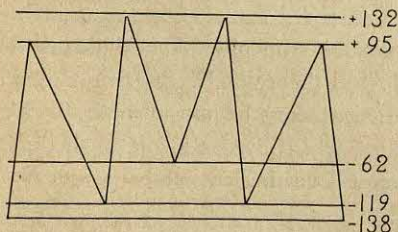


FIG. 2. HYDRAULIC EFFECT OF
 $2\cos 7x + \cos 10x$

After the three smallest hither-thither movements of the fluids lengthwise have been absorbed by the phragma; Tartini Pitch 4 is visible.

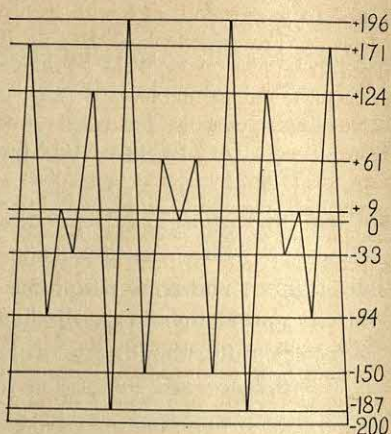


FIG. 3. HYDRAULIC EFFECT OF
 $\cos 7x + \cos 10x$
All the 10 maxima appear

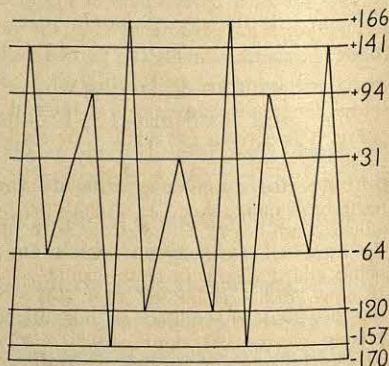


FIG. 4. THE ABSORPTIVE EFFECT ON
 $\cos 7x + \cos 10x$

After the three smallest hither-thither movements of the fluids lengthwise have been absorbed by the phragma only seven maxima appear.

I used audiogenerators, each with its own amplifier and its own loudspeaker; never telephones because telephone hearing is not normal hearing. In the beginning I tuned Tone 1 at 700~, Tone 2, as the musical lower fifth at 467~, and Tone 3, as the lower decima, at 280~.

Now, very gradually, I raised the frequency of Tone 2, thus bringing it slightly nearer Tone 1. At first, that makes the total sound of the combination 1 plus 2 very rough, but by degrees it becomes smoother again. When maximal smoothness is obtained, the clear Tartini Pitch 280 plus a weak and low rumble is heard. The middle tone is now identified as 490, having been raised from 467~. Thus the experimental result is that the ratio 10:7 (*i.e.* 700:490) produces the Tartini Pitch 4 (*i.e.* 4×70) plus a low rumble, but no difference tone and no summation tone.

My last statement must be qualified by taking into account the relative intensities of the sinusoids presented to the ear. With the ratio 10:7, this is observed: While 7 sounds forte or fortissimo, I introduce a very weak 10. At first 10 itself is then still inaudible although, with 7 being momentarily silenced, 10 is heard pianissimo. I put 7 back to fortissimo and very slowly continue to increase the intensity of the sinusoid 10. Suddenly the Tartini Pitch 4 appears *before Pitch 10 is heard even weakly*. Only somewhat later Pitches 10, 7, and 4 plus the rumble are all audible.¹

These results are made theoretically plausible by Figs. 1 and 2. In Fig. 1 the sinusoid 7 has the coefficient 2 and the sinusoid 10 the lesser coefficient 1. The *hydraulic* theory of cochlear mechanics is applied by first computing all the maxima and minima of the curve.² Fig. 1 demonstrates that under these relative intensities of the sinusoids Pitch 10 is *inaudible*, but Fig. 2, which results from the application of *absorption* (discovered in 1895) to the former figure, shows that the Pitch 4 must be heard in addition to Pitch 7 even under these conditions.³ Neither a difference nor a sum is derivable from Fig. 2.

Now we must have a look at Fig. 3 where the audibility of the higher tone, 10, is no longer impeded relative to the lower tone, 7. We count 10 maxima in the curve (in contrast to Fig. 1) in congruity with the experimental observation. Fig. 4 is derived from Fig. 3 by absorption of the three smallest (almost equal) hither-thither lengthwise movements of the scalar fluids.

Fig. 5 is derived from Fig. 4 by further absorption (± 79) of ordinate units. We

¹ I reported a similar experiment many years ago as follows: "The tuning Fork 5 must be sounded first, as strongly as possible, and it is necessary to have a fork which continues to sound strongly for quite a while. Then the bow is applied with the most delicate touch to Fork 8. One suddenly hears the low difference Tone 1 being added distinctly to the Tone 5 *although of Tone 8 no trace is yet audible*. If now Fork 8 is left to itself still vibrating and Fork 5 is stopped, the Tone 5 together with the subjective tone disappears and immediately one hears with surprising clearness Tone 8 which a moment ago was entirely inaudible" (M. F. Meyer, Introduction to the mechanics of the inner ear, *Univ. Mo. Stud., Sci. Ser.*, 2, 1907, 123). Nobody in the acoustical literature has ever given the slightest attention to this rather remarkable experience.

² M. F. Meyer, *How We Hear and How Tones Make Music*, 1950, 1-117. On the other hand Georg von Békésy, (Some electromechanical properties of the organ of Corti, *Ann. Otol. Lar.*, 63, 1954, 448-468) admits that his *vibratory* theory of the cochlea does not explain the facts of hearing.

³ Explained in my article, Theory for pitches 19, 15, and 11, plus a rumbling resulting from "sin 19X + sin 15X," *J. acoust. Soc. Amer.*, 27, 1955, 749-750.

count there exactly four maxima within the period. This agrees with the sensory observation. If we applied further absorption (I omit the resulting figure) we would explain the weak and low rumble. Neither a sum nor a difference of 7 and 10 would, however, come out; and neither were they heard. The above '79' is, of course, one-half of $(94 + 64)$ in Fig. 4.

Minor tritone. The above observations concerned the major tritone. We now describe the experiment with the minor tritone, 7:5. I last had the lower sinusoid tuned at 490~. I now raised it very gradually still further. At first this makes the total sound rougher. Then the sound becomes again smoother. When a maximal smoothness is obtained, the procedure is the same as with the major tritone. Playing both sinusoids forte, we hear in addition to them clearly the Tartini pitch identical with 300~ and a weak and low rumble.⁴

Fig. 7 is derived from Fig. 6 by absorption of the two smallest and equal hither-thither movements. Fig. 8 is derived from Fig. 7 by absorption of two

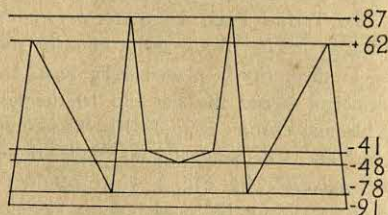


FIG. 5. HYDRAULIC EFFECT OF
 $\cos 7x + \cos 10x$
 After a further absorption of ± 79
 ordinate units; Pitch 4 is apparent.

more. In anatomical reality that means that in a certain region of the phragma the sensitive cells suffer only three charges per period, and the Tartini Pitch 3 must be heard.

As formerly with the major tritone, I now proceeded, letting the *intensity* of sinusoid 7 increase very gently *from zero* while that of sinusoid 5 was kept very great. Correspondingly we see in the legend of Fig. 9 the *different* coefficients of the combined sinusoids. Again, when I first notice the Tartini Pitch 3, not even a trace of Pitch 7 has as yet become audible. In Fig. 9 there are only 5 (not 7) maxima visible. Nevertheless we count in Fig. 10, derived from Fig. 9 by absorption, the

⁴ A colleague, Dr. C. A. Ruckmick, an expert in the field of audition, recently visited my laboratory and, after listening to the tones produced by the audiogenerators, made the following comments: "In earlier years I taught my students that the combination of two objective tones under certain conditions would produce a series of difference tones and possibly a summation tone, the latter of which was very much in doubt. All of them could, however, be calculated by mathematical formulae. I now regard that many of these so-called combinational tones can not be found under controlled experimental conditions and I suspect that they are merely mathematical constructs. When I listened to the sinusoids 700 and 500 in combination, I heard very clearly (perceptually, not imaginarily) a tone of 300 ~ and no tones of 200 and 1200 ~."

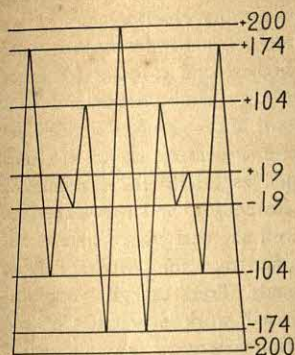


FIG. 6. HYDRAULIC EFFECT OF
 $\cos 5x + \cos 7x$

All the 7 maxima appear.

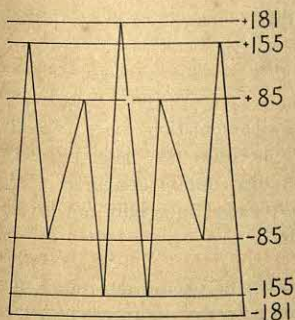


FIG. 7. HYDRAULIC EFFECT OF
 $\cos 5x + \cos 7x$

After the absorption of two
hither-thither movements by
the phragma.

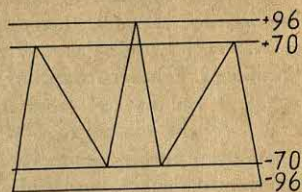


FIG. 8. HYDRAULIC EFFECT OF
 $\cos 5x + \cos 7x$

After a second absorption of
two hither-thither movements
by the phragma; Pitch 3 is
apparent.

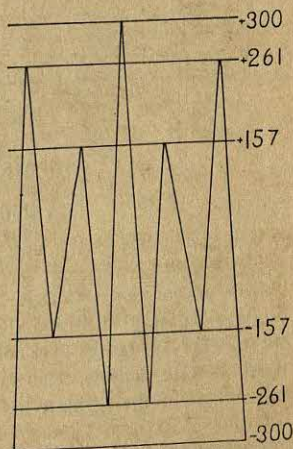


FIG. 9. HYDRAULIC EFFECT OF
 $2\cos 5x + \cos 7x$

The higher frequency is com-
pletely masked, but Fig. 10
will show that the Tartini
Pitch 3 must be heard never-
theless.

3 maxima demanded by hearing the Tartini Pitch 3. There is perfect agreement between the hydraulic theory and the above described sensory observation.

Naturally I made the same experiments also at lower frequency levels than those mentioned. I had no difficulty using 490~ as the highest and constant frequency where I had formerly used 700~, with 343~ for the major tritone and with 350~ for the minor tritone. When, however, I tried still lower absolute frequencies, the Tartini pitches became hard to identify. I have long suspected that the mechanical functioning of the cochlea is by Nature designed especially for a region of frequencies centered around 400~ approximately, extending about two octaves downward and two octaves upward. One might as well say that the cochlea is designed to work well for receiving the signals coming from the voice apparatus of mankind and similar species of mammalian animals. True, it will work also outside of those limits, but one must not expect that it will work in exactly the same way or as well. The accepted concepts of the evolution of species give to such a conclusion some plausibility.

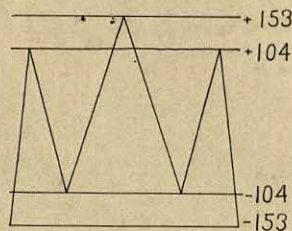


FIG. 10. HYDRAULIC EFFECT
OF $2\cos 5x + \cos 7x$

After absorption from Fig. 9
of the two smallest hither-
thither movements. A still
further absorption would ex-
plain the rumble heard.

The unconditional doctrine of difference tones and summation tones urgently ought to be eliminated and the truth be substituted. One must not, however, make the mistake of rashly overlooking or denying that under certain conditions 'beat tones' become audible. The theory of beat tones, neurological in origin, not in the mechanism of the cochlea, has been treated elsewhere.⁵

Curious third. Having gone so far with these experiments on the ratios 5:7 and 7:10, I decided it would be advisable to experiment on a further ratio in which the prime number 7 is one of the ratio-terms. Thus I chose the ratio 7:9. Since there is no conventional name for this interval, I call it the Curious Third. The following three lines show that it lies between the Major Third and the Fourth.

4:5 Major Third, equals 3863 mills of $1/12$ of $\log 2$.

7:9 Curious Third, equals 4351 mills of $1/12$ of $\log 2$.

3:4 Pure Fourth, equals 4980 mills of $1/12$ of $\log 2$.

I used a series of varying absolute frequencies. To facilitate comparison, I have

⁵ Meyer, Neurological theory of beat tones, *J. acoust. Soc. Amer.*, 28, 1956, 877-881.

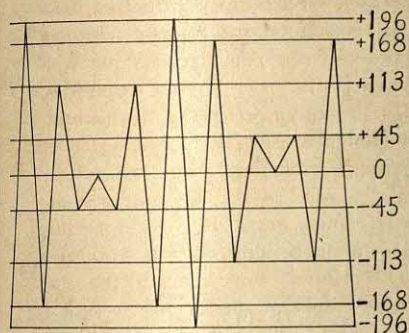


FIG. 11. THE COMPOUND CURVE
 $\sin 7x + \sin 9x$

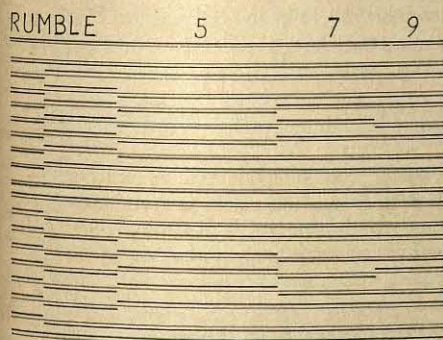


FIG. 12. SUCCESSIVE CONFIGURATIONS OF THE
 PHRAGMA

The beginning of the cochlea, at the round window, is the right edge of this figure. Going leftward, we go in the direction of the helicotrema. Numbers above the figure indicate relative pitch.

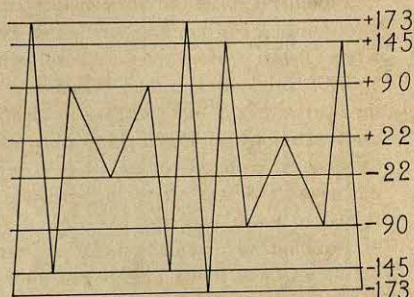


FIG. 13. THE HYDRAULIC EFFECT LEFT
 AFTER THE FIRST ABSORPTION OF EQUAL
 ORDINATE VALUES FROM FIG. 11

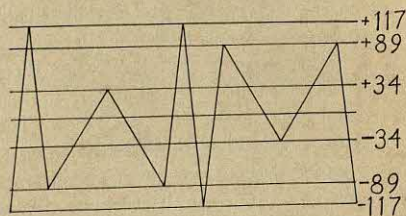


FIG. 14. THE HYDRAULIC EFFECT LEFT
 AFTER THE SECOND ABSORPTION FROM
 FIG. 11.

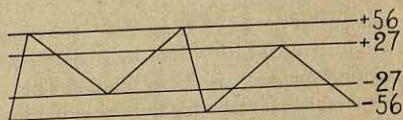


FIG. 15. THE HYDRAULIC EFFECT LEFT
 AFTER THE THIRD ABSORPTION FROM
 FIG. 11.

marked the frequencies by indicating the cycle units and also by the capital letters from A to H. The Tartini effects appear in parentheses for distinction from the sinusoidal airwaves. I believe that beat tones are unlikely to overlap the Tartini pitches because in my experience beat tones are very pronounced only with primary tones lying above 2000—. All subjective phenomena are thus likely to be Tartini pitches originating in the mechanism of the cochlea.

- (A) Unit 40, 360:280:(200 very weak but definite)
- (B) Unit 50, 450:350:(250 plus a weak rumble)
- (C) Unit 55, 495:385:(275 plus a rumble)
- (D) Unit 60, 540:420:(300 plus a rumble)
- (E) Unit 72, 648:504:(360 plus a rumble)
- (F) Unit 80, 720:560:(400 plus a rumble in the neighborhood of 80)
- (G) Unit 97, 873:679:(485 plus a rumble in the neighborhood of 97)
- (H) Unit 110, 990:770:(550 plus 110)

Two facts stand out as common to the Cases A to H: (1) The regular appearance of the Tartini Pitch 5 accompanying the primary tones 7 and 9 and (2) the absence of difference and summation tones.

Graphic illustration. I now use both of the graphic methods devised to illustrate my theory of hearing. They are two ways of demonstrating the hydraulic theory of the cochlear mechanisms. Both methods predict exactly the same sensory experiences. Fig. 12 uses the graphic method I devised in 1898. It shows the up and down deflections (jerks and temporary resting positions) that varying parts of the phragma assume according to my theory. Here the functioning part of the phragma is depicted in temporal order; the earliest configuration in the top line, the later configurations in the lines below. The bottom line, illustrating the natural end of the period, is identical with the first line. Figs. 13, 14, and 15 use the graphic method dating from 1896. They show in the abstract how, through mediation of the leathery ribbon of the cochlear phragma, the spatial absorption of ordinate values may result in simultaneous hearing of pitch 9, pitch 7, pitch 5 and pitch 3, or instead of a clear pitch 3 a rumble more or less resembling, perhaps vaguely, pitch 1.

Summation tone. As to a summation tone, it is enough to point out that in the whole acoustical literature nobody has ever claimed to have heard any *produced by the ear*. The not uncommon notion that Helmholtz has made such a claim is a misreading or mistranslation of his words. In the proceedings of the 1895 meeting of the Physical Section of the British Association for the Advancement of Science, the following report bearing upon the point at issue appears: "Monday's sitting was devoted to a discussion of the nature of combination tones. There was a general agreement that summation tones [made with tuning forks to prevent objective creation] have never been heard."⁶ This report prompted me to formulate the hydraulic theory of cochlear mechanics, starting from the exclusion of an audible summation tone.

Combining 4200~ with 5400~, I hear distinctly the *beat tone* 1200, thus 9:7(2), but no summation tone, and no Tartini pitches, that is, no 3000 and no 600; which means that Tartini pitches and beat tones are different in kind and origin—and that subjective summation tones are figments of the imagination.⁷

⁶ *Nature*, 52, 1895, 535a.

⁷ Meyer, *op. cit.*, *J. acoust. Soc. Amer.*, 28, 1956, 877-881.

THE INFLUENCE OF SIZE AND SHAPE ON THE DISCRIMINATION OF VISUAL INTENSITY

By MARVIN NACHMAN, University of Colorado

It has been demonstrated repeatedly that an inverse relationship exists between the size of a visual stimulus and the differential threshold of intensity. In demonstrating this relationship, most investigators have used circular stimuli and have suggested that the threshold is a function of area. Some of the proponents of the area-intensity view are Aubert,¹ Lasareff,² Heinz and Lippay,³ Cobb and Moss,⁴ Steinhardt,⁵ Holway and Hurich,⁶ and Crozier and Holway,⁷ who found the threshold to be inversely proportional to the area, or some exponential function of the area, of the stimulus. These theories which relate the difference limen to the area of the stimulus are necessarily limited, since it has been demonstrated that the shape of a figure will likewise affect its threshold. Thus, figures of equal areas but differing shapes (*e.g.* 10×10 vs. 20×5) result in different thresholds, a finding which can not be assimilated by any simple area-intensity relationship.

Two alternative formulations of the area-intensity relationship have been offered by Graham and his associates,⁸ and by Lamar, Hecht, Shlaer,

* Received for publication May 21, 1956. This study is based on a doctoral dissertation submitted in 1953 to the Graduate School of Boston University. It was conducted in the Psychology Section of the University's Optical Research Laboratory, under contract with the Photographic Reconnaissance Laboratory, Wright Air Development Center, USAF (Contract AF 33-038-15615). The author is indebted to Professors J. M. Harrison and L. J. Reyna for their assistance in this research.

¹ Hermann Aubert, *Physiologie der Netzhaut*, 1865, 82-89.

² P. Lasareff, Studien über das Weber-Fechnersche Gesetz: Einfluss der Grösse des Gesichtsfeldes auf den Schwellenwert der Gesichtsempfindung, *Arch. f. d. ges. Physiol.*, 142, 1911, 235-240.

³ Marianne Heinz and Franz Lippay, über die Beziehungen zwischen der Unterschiedsempfindlichkeit und der Zahl der erregten Sinneselemente, *Arch. f. d. ges. Physiol.*, 218, 1927, 437-447.

⁴ P. W. Cobb and F. K. Moss, The four variables of the visual threshold, *J. Franklin Inst.*, 205, 1928, 831-847.

⁵ Jacinto Steinhardt, Intensity discrimination in the human eye: I. The relation of $\Delta I/I$ to intensity, *J. gen. Physiol.*, 20, 1936, 185-209.

⁶ A. H. Holway and L. M. Hurvich, Visual differential sensitivity and retinal area, this JOURNAL, 51, 1938, 687-694.

⁷ W. J. Crozier and A. H. Holway, Theory and measurement of visual mechanism. III. ΔI as a function of area, intensity, and wave length, for monocular and binocular stimulation, *J. gen. Physiol.*, 23, 1939, 101-141.

⁸ C. H. Graham, R. H. Brown, and F. A. Mote, The relation of size of stimulus and intensity in the human eye: I. Intensity thresholds for white light, *J. exp.*

and Hendley.⁹ Both of these theories are equipped to handle the difficulties presented by the shape of the figure and the present experiment is an attempt to test them.

Graham's theory holds that if one considers the total area of a stimulus as composed of many unit-areas, then the contribution of each of these unit-areas to the differential threshold may be assumed to be a function of their distance from the center of the figure. The facts, that larger areas have lower thresholds and that shape affects the threshold are compatible with this assumption. In the above example, the 20×5 figure will be successfully predicted to have a higher threshold than the 10×10 because the unit-areas of which it is composed are further removed from the center.

The theory of Hecht and associates stresses the importance of boundary dimensions rather than area *per se* in determining thresholds. Using rectangles as their stimulus-figures, they found that they could successfully arrive at an equation to fit their threshold data by utilizing measures of perimeter (P) and a concept which they termed 'useful area' (UA). For background intensities of 17.5 ft.-L., the UA was empirically found to be the area within 1.5 min. of visual angle of the boundary of the stimulus, and the equation which best fit their data was $\Delta I/I = cP^{2/3}/UA$.

PROBLEM

The present experiments were designed to test the concept of useful area (UA) in several ways. In the test-figures used by Hecht and associates increases in the total area of the stimuli were always accompanied by increases in UA . A method of testing their equation is to vary the total area without changing UA or P . Thus, figures differing in total area, such as $10' \times 10'$ and $15' \times 5'$ would be predicted to have the same threshold since they are equal in P and UA . The UA is equal for these figures because it is defined as the area within 1.5' of the outside edge and they have equal outside edges. The same principle would apply for all figures not less than 3' in width. Below 3' in width, UA is the same as the total area. Experiment I of the present study utilized 12 rectangles in which UA and P remained constant while the total area varied, and 12 rectangles in which the UA and total area varied simultaneously. Experiment II was concerned with determining whether the same predictions could be made for more complex stimuli, *i.e.* figures composed of two rectangles. It also provided a test of Graham's theory

Psychol., 24, 1939, 555-573; C. H. Graham and N. R. Bartlett, The relation of size of stimulus and intensity in the human eye: III. The influence of area on foveal intensity discrimination, *J. exp. Psychol.*, 27, 1940, 149-159.

⁹ E. S. Lamar, Selig Hecht, Simon Shlaer, and C. D. Hendley, Size, shape, and contrast in detection of targets by daytime vision, I. Data and analytical description, *J. opt. Soc. Amer.*, 37, 1947, 531-545.

by varying the distance separating the two rectangles. The theory of Graham stresses the importance of the distance of unit-areas from the center of the figure and would predict that the DL would increase as two rectangles of uniform size were spatially displaced from each other.

METHOD

Apparatus. Fig. 1 is a schematic diagram of the apparatus, which consisted of two parts, a screen which provided the background intensity, and a long piece of 16-in. diameter cardboard tubing which housed the apparatus for presentation of the stimuli at various increments of intensity (ΔI).

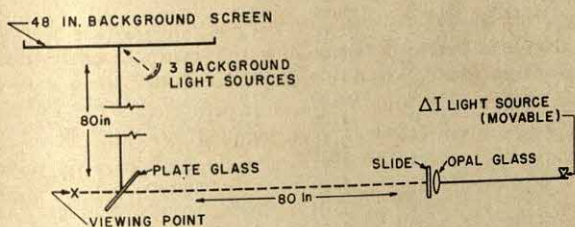


FIG. 1. SCHEMATIC DIAGRAM OF THE APPARATUS

The background screen was a 48-in. diameter disk of masonite, painted flat white and illuminated by three flood lamps. This screen was seen as a 30° field with an intensity of 17.5 ft.-L. at O 's station. Four small pieces of black tape were centered on the background screen to orient O and to insure foveal vision. The pieces of tape were $\frac{1}{8}$ -in. squares and were placed 2.5-in. apart (visual angle of $95'$) at the four corners of a diamond. The stimulus-figure always appeared at the center of these four orientation-dots.

A movable light-source was used to vary the intensity of the test-figures. The light was passed through four baffles, an opal glass, the test-figure (which was a prepared slide), an additional baffle, a $\frac{1}{2}$ -in. piece of plate glass, and finally through a 2-mm. artificial pupil to O 's eye. The plate glass, arranged at a 45° angle, transmitted the ΔI -light and reflected the background screen. The baffles were used to cut out stray light, and the entire cardboard tube was painted flat black. O viewed monocularly; an eyepatch over the other eye and a blackened headpiece further served to cut out stray light.

Stimuli: Experiment I. The stimuli in Experiment I were made from razor-blade edges attached to glass slides and so arranged as to permit the transmission of light in various sizes of rectangles. The slides were constructed under a microscope with a Filar Micrometer eyepiece, and it was found that the widths of the rectangles never varied more than 3 or 4 μ from the mean width. Experiment I used 24 slides: Slides 1-12, with a perimeter of $40'$, varied from $10' \times 10'$ to $19.9' \times 0.1'$; and Slides 13-24, with a perimeter of $80'$, varied from $20' \times 20'$ to $39.9' \times 0.1'$. Table I shows the dimensions of these slides.

Experiment II. Experiment II employed slides made up of two rectangles varying

in size and separation. Slide *A* was a $20' \times 20'$ square. Slides *B*, *C*, and *D* consisted of two rectangles of dimensions $20' \times 7'$, $20' \times 5'$, and $20' \times 3'$, separated 6', 10', and 14', respectively. Slides *E*, *F*, and *G* were each composed of two $20' \times 5'$ rectangles, separated by 3', 6', and 20', respectively.

TABLE I
DIMENSIONS OF THE SLIDES FOR SINGLE-FIGURE STIMULI USED IN EXPERIMENT I

Perimeter 40'				Perimeter 80'			
Slide no.	Visual angle (min.)	Useful area (sq. min.)	Area (sq. min.)	Slide no.	Visual angle (min.)	Useful area (sq. min.)	Area (sq. min.)
1	10.00 $\times$ 10.00	51.00	100.00	13	20.0 $\times$ 20.0	111.00	400.00
2	13.00 $\times$ 7.0	51.00	91.00	14	30.0 $\times$ 10.0	111.00	300.00
3	14.0 $\times$ 6.0	51.00	84.00	15	33.0 $\times$ 7.0	111.00	231.00
4	15.0 $\times$ 5.0	51.00	75.00	16	35.0 $\times$ 5.0	111.00	175.00
5	16.0 $\times$ 4.0	51.00	64.00	17	36.0 $\times$ 4.0	111.00	144.00
6	17.0 $\times$ 3.0	51.00	51.00	18	37.0 $\times$ 3.0	111.00	111.00
7	17.8 $\times$ 2.2	39.16	39.16	19	37.8 $\times$ 2.2	83.16	83.16
8	18.5 $\times$ 1.5	27.75	27.75	20	38.5 $\times$ 1.5	57.75	57.75
9	19.2 $\times$ 0.8	15.36	15.36	21	39.2 $\times$ 0.8	31.36	31.36
10	19.6 $\times$ 0.4	7.84	7.84	22	39.6 $\times$ 0.4	14.84	14.84
11	19.8 $\times$ 0.2	3.96	3.96	23	39.8 $\times$ 0.2	7.96	7.96
12	19.9 $\times$ 0.1	1.99	1.99	24	39.9 $\times$ 0.1	3.99	3.99

Observers. Two paid *O*s were used in Experiment I—*RDB*, a man (graduate student), and *LB*, a woman (undergraduate). The visual acuity (Snellen) for the preferred eye of each *O* was 20/20. *RDB* served as *O* in Experiment II.

Procedure. The serial order of the method of limits,¹⁰ was used to obtain a threshold for each slide. A trial consisted of presenting the ΔI light-source for 3 sec. by means of an interval-timer. The inter-trial interval was 7 sec., and the ΔI was then increased or decreased by 0.04 log intensity unit depending on *O*'s previous response. There were 65 such trials for each of Slides 1–24. The first 5 trials were systematically discarded as a warm-up period, and 10 of the trials were controls in which no stimulus was presented, to prevent *O* from responding indiscriminately. The remaining 50 trials were used to compute the threshold for each slide and, after thresholds for all 24 slides had been obtained, the entire procedure was repeated as a check on the reliability of the threshold measures. The final threshold measure was the mean of the 100 trials for each slide. The thresholds for Slides *A*–*G* in Experiment II were obtained in a similar fashion except that 35 trials were given at a time; 5 as a warm-up, 5 controls, and the remaining 25 to compute the threshold. The procedure was repeated 5 times giving a total of 125 trials on which to base the threshold. Experiment II was done 2 mo. after Experiment I.

RESULTS

Experiment I. Fig. 2 shows the mean log $\Delta I/I$ values of the two *O*s for each slide, as well as the curves predicted from the equation of Hecht

¹⁰ W. J. Dixon and A. M. Mood, A method for obtaining and analyzing sensitivity data, *J. Amer. stat. Assoc.*, 43, 1948, 109-124.

and associates. In arriving at the predicted curves, the value of the constant (c) was computed from their published threshold data. The value used was 0.13. A Lindquist trend-analysis was performed to determine whether the data for each O fit the predicted equations.¹¹ The analysis consisted of two tests: one for departure from pattern or shape of the predictions; and a second for the vertical displacement of the predictions. The F -values indicated that the equations predict the shape of the curve for both O s, although the equations and data are significantly different in vertical displacement. Differences in vertical displacement are a function

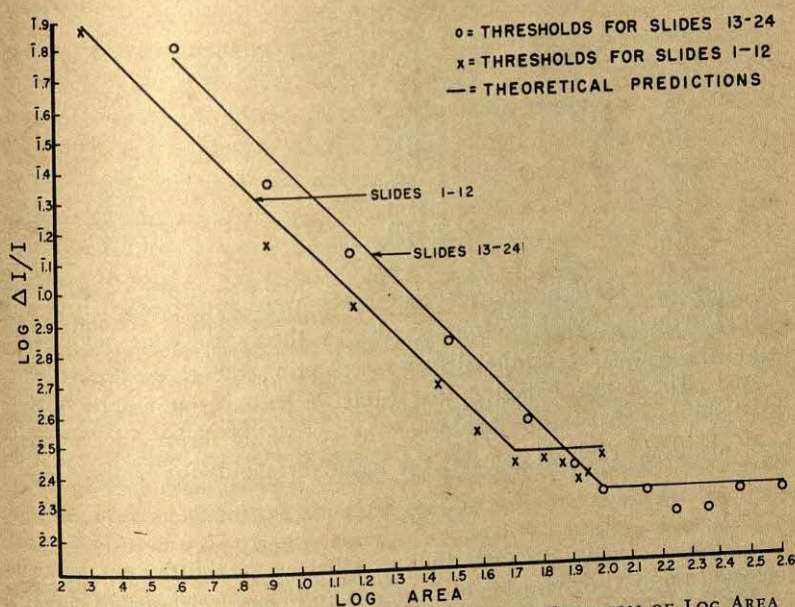


FIG. 2. DIFFERENTIAL THRESHOLD ($\text{LOG } \Delta I/I$) AS A FUNCTION OF LOG AREA

The solid lines are predicted from the equation of Lamar, Hecht, Shlaer, and Hendley.

of the value of the constant in the equation, and reflect differences in overall sensitivity. The differences between the overall thresholds of this study and those of Hecht and associates are probably most readily explained by the difference in criteria of 'seeing' used. The earlier study required O to recognize the position of a rectangle, which would undoubtedly result in higher thresholds than the present procedure of requiring O to report on presence or absence of the stimulus.

¹¹ E. F. Lindquist, Goodness of fit of trend differences, *Psychometrika*, 12, 1947, 65-77.

Experiment II. The dimensions of Slides *A*, *B*, *C*, and *D* were so selected as to predict approximately equal thresholds from the equation of Hecht and associates, despite the fact that they varied in total area. An *F*-test between the mean thresholds of these slides indicated that Slide *D* had a significantly higher threshold, but that the other three had very nearly the same thresholds, despite the fact that the total area of Slide *A* was about twice that of Slide *C*.

Slides *E*, *F*, *C*, and *G* each consisted of two rectangles of the same dimensions, but differed in the separation of the two rectangles. From

TABLE II
DIFFERENTIAL THRESHOLDS ($\log \Delta I/I$) FOR SLIDES A-G

Slide	Mean $\log \Delta I/I$	SD
A	2.4177	0.0241
B	2.4137	0.0531
C	2.4454	0.0652
D	2.5489	0.0425
E	2.5585	0.0426
F	2.4462	0.0664
G	2.4343	0.0610

the equation of Hecht and associates they were predicted to have the same thresholds, whereas from Graham's theory the thresholds were expected to increase as the distance separating the rectangles was increased. The *F*-test indicated that Slide *E* had a significantly higher threshold, but that the thresholds of the remaining three slides were almost identical. Table II contains the mean thresholds for Slides A-G.

DISCUSSION

Experiment I. Fig. 2 clearly indicates that as the width of a figure is increased up to 3' of visual angle, there is an inverse relationship between area and differential threshold (Slides 7-12 and 19-24). Beyond 3' of width, however, the boundary dimensions become important, and the concept of *UA* rather than total area seems to be paramount in determining the threshold. This was confirmed by Slides 1-6 and 13-18, which had approximately equal thresholds, even though Slide 1 had about twice the area of Slide 6, and Slide 13 had more than three and a half times the area of Slide 18. The data are interpreted as being contradictory to theories of visual discrimination which assume area of the stimulus to be a determinant of the threshold. The equation of Hecht and associates is consistent with the data, whereas the theory of Graham and associates seems inadequate to predict the break in the threshold curve beyond widths of 3'. From Graham's theory we would expect figures such as

$17' \times 3'$ and $10' \times 10'$ to have different thresholds, because the unit-areas of which they are composed differ in their distance from the center of the figure. The fact that Slides 1-6 and 13-18 did not differ in threshold appears to be contrary to Graham's hypothesis. It might be pointed out that, for figures of constant shape, increases in area do result in decreases in threshold, but in diminishing returns, which is also consistent with the *UA*-formulation. For example, Slide 13 ($20' \times 20'$) does have a lower threshold than Slide 1 ($10' \times 10'$).

A more careful examination of Fig. 2 indicates two further observations which may require some revision of the equation of Hecht and associates. The first concerns the level of the threshold for the two narrowest stimuli, $19.9' \times 0.1'$ and $39.9' \times 0.1'$. For both *O*s, these two slides have high thresholds compared to the trend of the other stimuli and, in fact, are responsible for a sizeable portion of the *F*-ratios of the trend-analysis. In the results of Hecht and associates, their equation satisfactorily fit all the data except two of the very small targets of areas less than 1 sq. min. The thresholds for these targets were too large and did not fit their equation, although this fact was not stressed. It may be concluded that the results of the present study support the suggestion that the equation of Hecht and associates is inadequate to predict the thresholds of exceedingly small stimuli. Limitations due to the size of cell-structures and neural connections undoubtedly become important factors at these small sizes.

The second observation concerns the point in the curve at which the threshold becomes constant as the area is increased, *i.e.* the exact definition of useful area. It is clear from the data that there is no continual decrease in threshold as area is increased, but the specific point at which the threshold becomes constant is not too apparent. It may be that the width of $3'$ is not critical, but that some other value may be more adequate. The data for one *O*, for example, show that slides $17' \times 3'$ and $37' \times 3'$ have slightly higher thresholds than the five other slides with which they should be equal, suggesting perhaps that a width greater than $3'$ would constitute *UA*.

Experiment II. Slides *A*, *B*, *C*, and *D* tested whether the threshold predictions made for one rectangle would hold also for stimuli composed of two rectangles. Slides *A*, *B*, and *C* appear to confirm the hypothesis inasmuch as they had almost equivalent thresholds. Slide *D*, consisting of two $20' \times 3'$ rectangles separated by $14'$, had a higher threshold. As will be seen below, the high threshold of Slide *D* cannot

be attributed to the large separation of the two rectangles and is more likely due to a decrease in useful area. It seems probable that widths greater than 3' can be useful, an assumption which was suggested by Experiment I.

The fact that the distance separating two rectangles does not seem to be a variable in determining the threshold is verified by Slides *F*, *C*, and *G*, which varied in separation distance from 5' to 20', but had almost identical thresholds. Slide *E*, consisting of the same two rectangles separated by 3', had a higher threshold. This finding again seems to confirm the notion that a width of 3' may be somewhat too small as the definition of *UA*. The data from Slides *E*, *F*, *C*, and *G* appear to contradict the expectations from Graham's theory, which states that the threshold will be influenced by distance of areas from the center of the figure. Separation within the narrowly defined limits tested, did not seem to be a factor.

It seems logical, from the foregoing, to conclude that thresholds are determined by boundary dimensions of a stimulus rather than area. It is noteworthy that contemporary work in physiology¹² and optics¹³ stresses the importance of boundaries or edges of stimuli, rather than area, in determining visual properties.

SUMMARY

The present experiments were designed to test a theory reported by Lamar, Hecht, Shlaer, and Hendley, which suggests that differential thresholds are determined by boundary dimensions of a stimulus rather than by its area. In Experiment I, the differential thresholds for 24 rectangles of different dimensions were obtained for each of two *Os*. Half of these rectangles were so selected that their boundary dimensions would lead to predictions of equal thresholds although their total area differed considerably. It was concluded that the equation of Hecht and associates adequately predicted the threshold data. Theories which relate the differential threshold to the area of the stimulus, and the theory of Graham and associates which suggests that the threshold is a function of the distance of unit-areas from the center of the figure, were found to be inadequate to predict the data. Experiment II involved testing and substantiating the same predictions for figures composed of two rectangles.

¹² W. H. Marshall and S. A. Talbot, Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity, in Heinrich Klüver (ed.) *Visual Mechanisms, Biol. Symposia*, 7, 1942, 117-164.

¹³ L. A. Jones and G. C. Higgins, Photographic granularity and graininess: IV. Visual acuity thresholds; Dynamic versus static assumptions, *J. opt. Soc. Amer.*, 38, 1948, 398-435.

THE EFFECT OF "RIGHT" IN A MODIFIED THORNDIKIAN SITUATION

By LYMAN W. PORTER, University of California

Postman and Adams have recently published experimental findings which they interpret as evidence that rewards function directly and automatically in accordance with Thorndike's Law of Effect.¹ In their experiment they used a modified form of a procedure devised earlier by Wallach and Henle.² The Ss are required to call out numbers from 1 to 10 each time a stimulus-word is presented, and their responses are immediately 'rewarded' (the number is called "right" by E) or 'punished' by being called "wrong" by E according to a pre-arranged key. This part of the procedure corresponds to a typical Thorndikian spread-of-effect experiment in which Ss are instructed to learn the correct word-number pairs. In the above experiments, however, the Ss are not instructed to learn; rather, they are told that the experiment is a study of extrasensory perception (ESP) and that therefore responses which are correct on one trial may or may not be correct on the next trial. Presumably this procedure provides a satisfying reward but nevertheless does not motivate the Ss to try to remember or to repeat correct responses on subsequent trials. Wallach and Henle found no evidence for the direct-action hypothesis whereas Postman and Adams, with a modified procedure, found positive results.

One important difference in the procedure of Postman and Adams was that they did not use the Wallach-Henle method of going through the list a number of times always rewarding new positions in the list without regard to the S's responses. This aspect of the Wallach-Henle procedure could have caused Ss to develop a set against repeating (since the same 'connection' was never 'right' twice in a row), and it could have accumulated proactive interference which minimized any tendencies to repeat. In the experiment of Postman and Adams, information about the correctness of response was given only on the first trial and its effect on repetition was tested

* Accepted for publication June 18, 1956. This paper is adapted from a doctoral dissertation submitted to Yale University. The author is indebted to Dr. Fred D. Sheffield, who directed the research, and to Drs. Neal E. Miller and William Kessen, for their assistance.

¹ Leo Postman and P. A. Adams, Performance variables in the experimental analysis of the law of effect, this JOURNAL, 67, 1954, 612-631.

² Hans Wallach and Mary Henle, An experimental analysis of the law of effect, *J. exp. Psychol.*, 28, 1941, 340-349.

only on the second trial, the third and last run through the list being used only to measure recall. These procedural improvements in the Postman-Adams experiment may account for their getting the effect which did not show up under the Wallach-Henle conditions.

The Postman-Adams procedure, however, involved several factors which might have spuriously increased repetition and recall of rewarded responses: (1) The same three serial positions—6, 13, and 20—were always rewarded, thus giving rewarded responses any possible advantages associated with those particular serial positions. (2) Rewarded words were kept in the same serial positions on the test-trial while the order of the punished words was scrambled, giving rewarded responses the possible advantage of a more distinctive cue (same location in sequence). (3) All Ss were given recall tests and later told the purpose of the experiment, thus revealing that the study was concerned with repetition and memory rather than with ESP. Communication to potential future Ss was therefore a real possibility. (4) Finally, only six particular words (three words each on two forms of the list) ever were rewarded. These six words also were used in the 'punished' positions, but 15 of the 'punished' words analyzed were never used in the 'rewarded' positions. With this lack of balance and such a small sample of rewarded words, there is a possibility that these six particular words might by chance have stronger initial word-number associations than the always-punished words.

Another way in which the Postman-Adams procedure differed from that of Wallach-Henle was that, whereas the latter presented their words at a rate comparable to Thorndike's (2.2 sec.), Postman and Adams used a rate of approximately 5 sec. per word. There is thus the possibility that the positive finding by Postman and Adams is due to the use of a slower rate of presentation. A theoretical basis for such a possibility can be suggested without recourse to Thorndike's hypothesis. It is possible that saying 'right' in a typical Thorndikian experiment does not directly stamp in the responses as Thorndike hypothesizes, but rather acts to motivate active implicit rehearsal which produces learning by a principle of contiguity or frequency. A reduced form of the same phenomenon might take place in Wallach and Henle's situation. For example, S might review implicitly the 'right' word-number pairs, wondering what he did on those occasions to make ESP 'work'; or, all Ss might have general tendencies to review implicitly 'successes' of any kind. Any such tendencies would give differential implicit rehearsal of rewarded responses, even if Ss were not motivated to recall or learn them for later use. If such implicit rehearsal were actually producing learning by contiguity alone, there would be more time for the process at the Postman-Adams rate than at the Wallach-Henle rate. If the Postman-Adams phenomenon disappeared at the Wallach-Henle rate, there would be little support for direct action in their experiment.

The purpose of the present experiment was twofold: (a) to repeat the Postman-Adams study at approximately their rate of presentation but with the introduction of controls for the previously mentioned procedural artifacts, and (b) to use the same modified procedure at approximately the Wallach-Henle rate.

METHOD

Subjects. The Ss were 64 Yale undergraduates enrolled in a course in introductory psychology.

Procedure. The 64 Ss were divided into two groups of 32 each. One group received the stimulus-words at approximately the Postman-Adams rate (changed from 5 sec. to 6 sec. because the latter rate permitted the use of a standard presentation-apparatus). This group is hereafter called the 6-sec. group. The other group received the stimulus-words at approximately the Wallach-Henle rate (2 sec.). This group will be referred to as the 2-sec. group.

The Postman-Adams procedure was followed as closely as possible except for the control modifications mentioned above. The stimulus-objects were the 25 three-letter words used by Postman and Adams, together with three additional common three-letter words—28 words rather than 25 words were needed for balancing purposes. These words were presented with a memory drum, and S was instructed to respond to each with a number from 1 to 10. In the 2-sec. group, a word was shown for 1.9 sec. every 2 sec.; in the 6-sec. group, a word was shown for 1.9 sec. every 6 sec.

As part of the balancing procedure used, positions 5-8, 13-16, and 21-24 were designated beforehand as key-positions to be analyzed. All responses except three were called 'wrong' by E the first time through the list. These three responses were called 'right,' one right being given in each of the three key-blocks of four words each (5-8, 13-16 and 21-24). Within each experimental group, eight Ss had the first word in each key-block rewarded, eight had the second word in each block rewarded, and so forth. Since the rewards came only from the key-positions, only 'wrongs' that also came from these positions were analyzed. Thus, for each S three 'rights' and nine 'wrongs' were analyzed in the results. For half the Ss, one set of 12 words was used in the key-positions; for the other half, an entirely different set of 12 words was used. For 24 of the 28 words, therefore, this procedure balanced both for specific words and for specific serial positions of 'rewards' and 'punishments.' As in the Postman-Adams procedure, the first two and the last two words were not considered. For each S the list was repeated a second time, with all the words except the first two and the last two (which were interchanged) scrambled in random fashion. There was no 'reward' or 'punishment' on this second trial.

When S entered the experimental room he was given the following instructions to read:

You have undoubtedly heard of mental telepathy and other forms of extrasensory perception. Subjects are asked to attempt to perceive items which they cannot possibly be aware of by any of the known physical senses. A large number of experiments have reported results which have been interpreted as evidence that this ability exists. We are investigating the influence on this ability of such factors as knowledge of results, rate of presentation of stimuli, and presence of distraction while you attempt the perception.

We have prepared a list of pairs of words and numbers. The words will be presented mechanically, and your task is to attempt to select correctly the numbers that go with them. We have used numbers from 1 through 10, so you can choose from 10 numbers in each case, and with this large selection you will not be correct very often by chance. After each judgment I will tell you if you were right or wrong.

Thinking about your performance during the experiment definitely will not help you, so just try to give your responses as *quickly as you can*. Your first impression is supposed to give you the best score. Don't mind if we hurry you [go slowly] through the experiment; the rate of presentation of words is one of the variables we are studying.

We have a list of 28 words, and this list will be presented more than once, but the words will not remain in the same order. Also, the numbers assigned to the words will change in a random fashion from one presentation to the next. Thus a

number which is correct on the first presentation may or may not be correct on the next.

Are there any questions?

After these instructions were read, they were discussed informally by *E* to make sure that *S* had understood what was to happen and what his task was to be. Then *E* said:

In order to give you practice at responding at the rate at which the words will appear, I am going to have you call out numbers from 1 to 10 at random to three or four dashes—like the one you see in the window—that will appear at the rate that the words will in the actual experiment. This is just to get you used to the rate at which the words will come in the experiment and to be sure we understand the instructions.

S was then given this practice of calling out a number to each of four dashes which were presented at the rate to be used in the experiment. At the conclusion of this practice, *S* was told:

Now we'll begin the actual experiment. Remember, you are to try to select or guess the correct numbers from 1 to 10 that go with the words. I'll tell you after each guess as to whether you are right or wrong.

The first list was then presented, *S* calling out a number to each of the words and *E* responding with 'right' or 'wrong.' At the conclusion of the first list, *E* orally gave the *S* the following instructions:

Now we will go through the words again, using a different order of the same words that I gave you last time. This time you will not be informed as to whether you are right or wrong; that is, I will remain silent throughout. Remember, also, that this time the words and numbers have been re-paired by a random procedure; thus the number paired with a particular word this time may or may not be the same number as last time.

The scrambled list was then presented, with *E* saying nothing after each response. At the conclusion of the second list *E* talked briefly with *S* about *ESP* and his performance from this standpoint. He also said he would appreciate it if *S* would not discuss the experiment with others. It was *E*'s impression that he got a uniformly good response to this request, which was considered important since one of the major control-features of this experiment was the attempt not to divulge anything concerning its purpose to any *S*, either during the running or afterward.

RESULTS

The percentages of 'right' and 'wrong' numbers repeated in the appropriate positions (*i.e.* to the same words) on the second trial are presented in Table I. These results are calculated for the 12 key-positions (5-8, 13-16, and 21-24) in which 'rewards' and 'punishments' are balanced for specific words and specific serial positions. Table I shows that in both groups the responses called 'right' were repeated more often than responses called 'wrong,' the differences in each case being significant at the 7% level of confidence with a one-tailed *t*-test. Combining the 2-sec. and

6-sec. groups gives a t of 2.16 which is significant at the 2% level of confidence.³

The first purpose of the present experiment was to try to verify the results of Postman and Adams using approximately their rate of presentation, but incorporating additional controls that were felt to be necessary. The 6-sec. group in this experiment tended to confirm the Postman-Adams results, but the effect was not as large as in their experiment. The difference is primarily reflected in the fact that the percentage of 'rewarded' responses repeated in their experiment was 26.7, whereas in the present experiment it was only 19.8. The percentages of repetition for 'wrong'

TABLE I

PERCENTAGE OF 'RIGHT' AND 'WRONG' RESPONSES REPEATED ON THE SECOND TRIAL			
Responses	6-sec.	2-sec.	Combined groups
Right	19.8	16.7	18.2
Wrong	14.2	11.1	12.6
Diff.	5.6	5.6	5.6
df.	31	31	62
t	1.49	1.56	2.16
p (one tail)	0.07	0.07	0.02

responses in the two experiments were nearly equal—13.3 for Postman and Adams and 14.2 in this experiment. The failure here to obtain as high a level of repetition of 'right' responses may perhaps be attributed to the various control factors introduced in the present experiment, although it is impossible to assign the difference to any particular control or group of controls. There is also the possibility that the manner in which the stimulus-words were presented (*i.e.* mechanically by a memory drum at an unvarying rate) may somehow have tended to reduce the effects of the 'right.' The Ss may, for example, have paid less attention to the words than in the Postman-Adams procedure where the words were presented by E on index-cards.

A second major purpose of this experiment was to compare the performance of the group studied at the Wallach-Henle rate (2 sec.) with the 6-sec. group. The results showed that although the 2-sec. group did not quite produce a significant effect, it tended about as strongly in that direc-

³ A one-tailed t -test is used because the direction is implied by the direct-action theory being tested here. Although the frequency-distributions for percentage of 'right' and 'wrong' responses repeated are highly skewed in a positive direction (because of the high frequency of zero-values), the distribution of differences (number 'right' minus number 'wrong') is fairly symmetrical about a modal value, and the use of the t -test appears justified.

tion as did the 6-sec. group. Thus, the Postman-Adams phenomenon was not eliminated when the experiment was carried out at the Wallach-Henle rate. The 'wrong' responses at the 2-sec. rate in the present experiment were repeated about as often as in the Wallach-Henle experiments, but the repetition of 'right' responses was about 6% higher. These findings suggest that the original controls introduced by Postman and Adams—*i.e.* repeating the list only twice and not saying "right" or "wrong" the second time—rather than the change in rate of presentation were responsible for the difference between their results and those of Wallach and Henle.

While the 2-sec. group did not differ from the 6-sec. group with respect to the difference between 'right' and 'wrong' responses repeated, its overall level of repetition was lower. A 2×2 table (using 0-2 and 3-or-more as the two frequency categories) gave a chi-squared significant at the 6% level (two-tails).

One possible explanation for the present results, and those of Postman and Adams, is the factor of perceptual-isolation described by Von Restorff.⁴ Thus, saying 'right' only three times in a long list may make those particular pairs better learned because of isolation or vividness. Postman and Adams considered this possibility and conducted a second study in which they varied the degree of isolation,⁵ one group with 3 'right' and 18 'wrong' responses, another group with 10 'right' and 11 'wrong' and a third with group 18 'right' and 3 'wrong.' They concluded that isolation was not the chief factor because all three groups showed more repetition of 'right' than of 'wrong' responses. Their experiment was not entirely convincing, however, partly because the difference in frequency of repetition was very small for the group with 10 'right' and partly because no mention was made of how the group with 18 'right' interpreted an ESP situation in which 'chance' was 10% but in which they were 'right' about 85% of the time. It should be noted in any case that isolation or vividness can be reduced to differential rehearsal: anything that causes *S* to think more about a particular pair would tend to increase its likelihood of being repeated or recalled.

Subsequent to the completion of this experiment, Bitterman presented the results of an experiment which suggests another possible explanation of the Postman-Adams phenomenon in terms of differential information

⁴ Hedwig von Restorff, Ueber die Wirkung von Bereichsbildungen im Spurenfeld, *Psychol. Forsch.*, 18, 1933, 299-342.

⁵ Postman and Adams, 'Isolation' and the law of effect, this JOURNAL, 68, 1955, 96-105.

supplied by 'rewards' and 'punishments'.⁶ In his study he used three groups: (1) a group in which no responses were called 'right' or 'wrong'; (2) a group like that of Postman and Adams; and (3) a corrected group in which announcements of 'wrong' were followed by specification of the 'correct' number. He found that the effect of this correction was to reduce the frequency of repeating 'wrong' responses and to increase the frequency of repetition of the 'correct' numbers supplied by *E*—numbers which were as irrelevant as the 'wrong' numbers to the *S*'s task the second time through the list. Bitterman interprets this finding as evidence that it is not the satisfying effect of the reward, but the information as to what number is 'correct' provided by 'right' that accounts for the increased repetition of rewarded responses in the Postman-Adams study. He suggests that the *ESP* procedure may not entirely eliminate intent to learn, or that "incidental learning may be structured by the information in question."

Although the results in Bitterman's group, which was most like the Postman-Adams group, were in the same direction as those of Postman and Adams, it is worth noting that the small difference (3.2%) between 'right' and 'wrong' responses did not approach statistical significance, despite an *N* (=50) two and one-half times larger than the *N* (=20) of Postman and Adams. As in the present experiment, Bitterman never informed his *Ss* about the true purpose of the work and used a procedure in which the words analyzed were balanced for frequency of use in rewarded and punished positions. Bitterman's lack of significant results with this group therefore lends further support to the possibility that the reason for not getting a larger effect in the present experiment was the introduction of the procedural controls. Postman and Adams, without these controls, found a 13.3% difference in repetition of 'right' vs. 'wrong' responses, the present experiment, with these controls, showed a 5.6% difference, and Bitterman, with comparable controls, a 3.2% difference.

The procedural controls cannot, however, account for the entire phenomenon, since the combined 2-sec. and 6-sec. groups (*N* = 64) of the present experiment showed a significant effect. The comparison of the 2-sec. and 6-sec. groups did not lend support to the hypothesis that the Postman-Adams phenomenon was a result of the differential implicit review of the 'right' responses. The fact that a difference between the two groups did not appear implies that either: (1) implicit rehearsal is not an important factor in producing a differential effect of 'right' vs. 'wrong'; or

⁶ M. E. Bitterman, Information and effect in incidental learning, this JOURNAL, 69, 1956, 410-416.

(2) that even at a 2-sec. rate of presentation a substantial amount of implicit rehearsal takes place.

SUMMARY

This experiment was designed to measure the effects of 'right' and 'wrong' on the repetition of responses in a Thorndikian word-number situation under *ESP* conditions. It was also designed to reconcile conflicting results in earlier studies on this problem by Wallach and Henle and by Postman and Adams. Controls for possible procedural artifacts of these experiments were introduced and the rate or presentation of stimulus-words was varied to investigate that factor as a possible major source of the discrepancy in the previous results.

In the present experiment, responses called 'right' were repeated significantly more often than responses called 'wrong,' despite the introduction of the procedural controls. The magnitude of this difference was the same for both a 2-sec. and a 6-sec. rate of presentation, indicating that differences in rates of presentation probably cannot account for discrepancies in the results of previous experiments.

AN ANALYSIS OF INDIVIDUAL DIFFERENCES IN OLFACTORY THRESHOLDS

By F. NOWELL JONES, University of California, Los Angeles

Theories of the action of odorants on olfactory receptors have thus far failed to carry conviction.¹ In other words, the relevant physical-chemical dimensions of the olfactory stimulus are not known. The present study was designed to determine the way in which the human organism classifies odors, using as the method of determination the analysis of individual differences in olfactory thresholds. This approach assumes that (1) consistent and systematic threshold-differences exist, and (2) dimensions along which individuals differ are meaningful in regard to underlying function.

Although the general approach adopted here is commonplace in the field of testing, it has had very restricted application to psychophysical problems, perhaps because the relatively mature status of research in many sense-departments has not warranted the use of exploratory methods. A desire to test the method rather than to contribute unique information dictated the application of factor analysis to vision.² In olfaction the situation is different, and exploratory methods are entirely justified. Some factor analyses of olfactory data have been reported, but they have been concerned with judgments of esthetic value, not with measures of sensitivity.³ Closely related to the factor analytical approach is the analysis of partial anosmias.⁴ The analogy to the analysis of color blindness to discover basic visual processes is evident. Potentially fruitful though this approach may be, it has only begun to suggest possible basic classes of odor, and much more exploration of the dimensionality of odors will be necessary before a clear pattern emerges.

The present study involved the factor analysis of absolute olfactory thresholds for 20 substances and 84 Os. Factor analysis was chosen as the statistical technique

* Received for publication June 6, 1956. The work reported here was carried out in part under Contract N6 onr-27515 between the University of California and the Office of Naval Research, and in part through the support of the Committee on Research, University of California, Los Angeles. Several research assistants participated in this work at various stages: Slade Hulburt, J. H. Standing, William Bame, and J. W. McKelligott were all of very great help.

¹ F. N. Jones and M. H. Jones, Modern theories of olfaction: A critical review, *J. Psychol.*, 36, 1953, 207-241; Carl Pfaffmann, Taste and smell, *Ann. Rev. Psychol.*, 7, 1956, 391-408.

² Jones and Jones, A second factor analysis of visibility data, this JOURNAL, 63, 1950, 206-213.

³ J. G. Beebe-Center, General affective value, *Psychol. Rev.*, 36, 1929, 472-480; William Stephenson, The inverted factor technique, *Brit. J. Psychol.*, 26, 1936, 344-361; M. D. Eysenck, An experimental and statistical study of odor preferences, *J. exp. Psychol.*, 34, 1944, 246-252; E. H. Hsü, A factorial analysis of olfaction, *Psychometrika*, 11, 1946, 31-42.

⁴ M. Guillot, Anosmies partielles et odeurs fondamentales, *C. R. Acad. Sci.*, 226, 1948, 1307-1309.

since it represents a convenient method of finding structure in a set of results based upon systematic individual differences. Even though the end-result after rotation might not represent the most meaningful dimensions, the possibility of finding structure at all would suggest the systematic character of the differences found. Absolute thresholds were chosen because of their relative simplicity. One might work at some other level of odorous intensity, but considerable preliminary work would be required to establish equated scales of intensity for all of the odors used.

Apparatus. The apparatus has been described in detail elsewhere.⁵ In brief, it permits injecting into *O*'s nostril a 40-cc. 'blast' of air with known odor-content. The blast-duration was set at 0.8 sec. and only odor-content varied from blast to

TABLE I
CORRELATION MATRIX*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																				
2	68																			
3	48	68																		
4	57	68	50																	
5	65	63	42	52																
6	36	56	51	59	45															
7	34	40	36	41	44	48														
8	31	36	35	26	31	47	30													
9	48	46	32	51	42	42	13	39												
10	41	54	43	54	55	46	40	19	51											
11	50	57	53	68	51	62	46	35	52	51										
12	33	30	23	43	39	42	54	30	30	40	37									
13	28	34	33	08	26	32	28	23	33	34	31	15								
14	46	43	38	40	39	38	36	19	43	35	51	31	35							
15	43	53	50	47	51	43	37	36	48	56	58	28	30	47						
16	30	37	33	43	24	42	20	56	46	21	36	27	21	18	23					
17	37	40	25	47	45	42	29	35	52	36	46	39	28	24	38	44				
18	25	42	42	36	42	36	22	19	32	40	45	14	18	20	36	21	48			
19	52	44	38	49	29	35	13	27	40	26	46	19	31	36	20	48	43	27		
20	47	66	70	52	42	42	29	28	41	50	55	28	36	32	47	40	45	54	41	

* Substances are numbered as in Table II. Decimal points in the body of the table are omitted.

blast. The blast-temperature was held at 26°C. During the threshold-determinations, *O* sat in an air-conditioned glass-lined booth kept at $70 \pm 2^\circ\text{F}$. The air in the booth was in constant recirculation through an activated charcoal filter. It is not claimed that this booth represented a *camera inodora*, but rather that the ambient odor was held to a low and constant level.⁶

Observers. In all, there were more than 90 *O*s, but only 84 of these are included in the analysis. The others were eliminated for various reasons, mostly because of failure of the thermostat in the constant-temperature bath. Two were eliminated because all thresholds were extremely low and equal on the syringe-scale, suggesting some error of observation. Nearly all the *O*s were undergraduate or graduate majors in psychology, while a few were staff-members. Both sexes were represented in the group.

Substances. With due consideration of the time required to obtain thresholds, it was decided that 20 substances would represent a fair test of the question of

⁵ F. N. Jones, An olfactometer permitting stimulus specification in molar terms, this JOURNAL, 67, 1954, 147-151.

⁶ The Pur Air Division, Barnebey-Cheney Corporation, very kindly provided a room air purifier which served to keep odor levels down in *E*'s space. I am indebted to Mr. Theodore Barnebey for this kindness.

whether systematic differences exist. The final choice of these substances was made on the basis of several criteria: (1) availability in pure form⁷; (2) adaptability to the apparatus in its present form, *i.e.* liquidity at 26°C; (3) availability of the data necessary to calculate vapor-pressure, hence concentration; (4) inclusion of a few series, *e.g.* the different butanols; and, (5) inclusion of as large a variety of pure substances as possible, including some consideration of the Guillot, Le Magnen, and Zwaardemaker classifications.⁸ The 20 substances finally chosen are listed in Table I and their sources of supply indicated. In general, these substances are fairly simple molecularly. Comparisons of substances differing in chain-length, shape, and side chain are possible.

TABLE II
FACTOR-LOADINGS BEFORE AND AFTER ROTATION

Substance*	Before rotation					After rotation				
	I	II	III	IV	h ²	G	I	II	III	IV
1 n-Butanol	70	12	08	-12	52	45	03	20	30	02
2 iso-Butanol	80	24	-14	-12	73	48	03	46	21	-12
3 sec-Butanol	69	24	-32	06	68	38	04	58	-06	-10
4 tert-Butanol	76	14	-17	-19	66	49	14	42	24	-16
5 n-Propanol	70	21	23	-18	63	43	-09	12	43	03
6 Allyl Alcohol	72	-10	-12	-14	56	52	31	24	22	-02
7 Ethylene Chloride	55	08	21	-31	45	38	03	03	49	-06
8 n-Butyric Acid	52	-40	-08	-19	47	47	53	-01	24	03
9 n-Caprylic Acid	66	-22	15	24	56	47	22	-05	03	44
10 n-Heptane	67	19	20	08	52	39	-12	13	19	21
11 n-Octane	79	08	-04	07	63	50	09	30	10	13
12 Ethyl Acetate	52	-10	29	-31	46	41	15	-14	52	04
13 Amyl Acetate	44	-11	18	29	38	35	-04	03	-04	42
14 Benzene	57	07	21	10	38	37	-14	18	13	22
15 Cyclohexane	67	22	16	13	54	49	67	09	05	05
16 Pyridine	54	-50	-27	-08	62	49	32	-03	08	32
17 Eugenol	62	-28	07	12	49	47	00	33	-17	16
18 Saffrole	53	12	-14	26	39	29	00	33	-05	13
19 Benzaldehyde	56	-19	-18	10	39	41	33	20	-05	13
20 Methyl Salicylate	72	15	-25	24	66	41	06	48	-16	10

Transformation Matrix				
I	II	III	IV	
I	205	279	228	158
II	-880	540	-006	-344
III	-374	-794	523	549
IV	-208	014	-821	745

* 1-11 and 14-16 from Eastman, 12 from Baker, 13 from Mefford, 17-20 from Fritzsche Bros.

Procedure. The choice of procedure presented a practical problem. The central objective of the research was to obtain sufficient information for factor analysis. This goal was best served by having a large number of Os. It was decided, therefore, to simplify the procedure of measuring the threshold to permit the use of more Os, and also to permit the complete testing of each O in but one sitting, which ruled out day-to-day changes in threshold. Each O spent from 2-3 hr. in the experimental setting. This time-interval permitted but one ascending series for the determination of each threshold, a procedure justified by (1) the 1000 to 1 range

⁷ Suggestions by Dr. Ernest Guenther, Vice-President of Fritzsche Bros., Inc., were of great help in this regard.

⁸ Guillot, *op. cit.*, 1307-1309; J. Le Magnen, *Analyse d'odeurs complexes et homologues par fatigue*, *C. R. Acad. Sci.*, 226, 1948, 753-754; Hendrik Zwaardemaker, *Die Physiologie des Geruchs*, 1895, 207-238.

of individual differences found, and (2) the presence of systematic differences in the data.

For each *O*, the order of substances was random. The first threshold-determination was always for citral, which was used solely as a practice substance. Recognition-thresholds were used to minimize the effect of the sensing of the blast itself. This procedure required that *O* be given a whiff of the odor before the threshold for each new substance was determined. Changing the apparatus to the new substance required almost exactly 2 min., and that was the time which elapsed between the sample whiff and the beginning of threshold-determination. Trials in a given series were spaced by about 20 sec., a value chosen on the basis of other experimentation.⁹ The steps of concentration were arranged to approximate an increase of 30% per step.

Result. Pearson product-moment correlation coefficients were calculated for each pair of substances. The resulting matrix is shown in Table I. In grouping the data in 15 intervals for the calculations, the intervals were set according to a constant ratio for each substance, the resulting distributions appearing much less skewed than the original distributions. The correlation matrix was next factor-analyzed by the centroid method to an end-point which retained one factor beyond that required by Tucker's criterion, or four factors.¹⁰ The unrotated factor-loadings are shown in Table II.

Since there was some possibility that all four factors did not represent actual test-variance, rotations were carried out by different methods and with two, three, or four factors retained. The first, orthogonal solution has been reported elsewhere.¹¹ The solution finally adopted was obtained by Thurstone's analytical method, with subsequent inspection of the factor plots.¹² The resulting oblique factor-loadings and the transformation-matrix are given in Table II, along with a general factor (*G*) obtained according to Thurstone's Case 4.¹³ Four factors were retained because, first, every hyperplane is satisfactorily determined, and, second, high loadings remain on each.

Discussion. The general factor may be ascribed either to differences in over-all sensitivity from *O* to *O*, or to differences among *O*s in the criterion of threshold. Although it is not possible from the data to decide which of these possibilities is correct, it is plausible to assume that both were in some measure operating. Certainly, experience with other senses suggests that individual differences in sensitivity should be found. The bases of these differences are not clearly known.

Before discussing the various group factors, it is well to consider what meaning they may be expected to have. Within the limits of errors of measurement, the factor-loadings represent the amount of common variance which the thresholds for a

⁹ F. N. Jones, The reliability of olfactory thresholds obtained by sniffing, this JOURNAL, 68, 1955, 289-290.

¹⁰ L. Tucker, reported in L. L. Thurstone, Primary mental abilities, *Psychomet. Monogr.*, 1, 1938, 66-67.

¹¹ F. N. Jones, An analysis of individual differences in olfactory thresholds, *Amer. Psychol.*, 9, 1954, 400-401 (Abstract).

¹² Thurstone, An analytical method for simple structure, *Psychometrika*, 19, 1954, 173-182.

¹³ Thurstone, *Multiple Factor Analysis*, 1947, 426-427.

given substance have with the other thresholds in this particular sample of substances and *O*s. Unless the choice of substances has been very fortunate indeed, we cannot expect the whole of 'odor-space' to be mapped by the one experiment. Variance not accounted for may represent unreliability, specific factors, or lack of sampling of a given region of the odor-space. The most highly loaded substance on a given factor may not, therefore, represent the 'type-substance' for the given dimension, but may merely represent the best example of the given dimension in the sample under consideration.

The interpretation of the first-order factors involves the consideration of both chemical and psychological dimensions.¹⁴ Factor I is most strongly represented by pyridine and n-butyric acid. There seems to be no straightforward chemical reason for this grouping, and when it is considered that eugenol and benzaldehyde have appreciable loadings on this factor the grouping does not seem to be one merely of unpleasantness. Inspection of a Harris-type plot of the original unrotated loadings suggested that eugenol lies on a line extending toward ethyl acetate while benzaldehyde tends toward the isomeric butanols.¹⁵ Factor II is best represented by the isomeric butanols. Harris-plots indicated that safrole and methyl salicylate both tend toward amyl acetate. Another substance which has its highest loading on this factor is n-octane. Factor III is characterized by high loadings on ethyl acetate and ethylene chloride. Except for sec-butanol, the aliphatic alcohols also are represented by loadings on this factor. All of these substances have rather low molecular volumes, but pyridine, which also is low, has no loading on this factor. Factor IV is best exemplified by n-caprylic acid and amyl acetate. Both benzene and cyclohexane have loadings on this factor, suggesting that they have some stimulating property in common. In general, the substances with significant loadings on Factor IV are quite unlike the various alcohols, none of which show appreciable loadings on this factor.

Since the factors do not clearly represent either chemical dimensions or homogeneous odor-qualities, their interpretation is difficult, but acceptable hyperplanes were obtained, and it is therefore necessary to consider what this fact implies. First of all, it appears that the odor-space has not been completely explored. One might anticipate that the addition of more substances would increase the number of factors, perhaps materially. Secondly, the general factor is of appreciable magnitude. The implications of this result have been considered previously. The third point requires that we assume tentatively that the basic stimulating process for all of the substances in this sample is the same. This is not an unreasonable assumption considering the particular substances used and the arguments concerning the meaning of the group-thresholds considered elsewhere.¹⁶ Now most research has shown that the variety of receptors must be rather large—certainly no one has unequivocally shown a limited number of primaries.¹⁷ The factor-loadings would, by

¹⁴ Chemical dimensions considered were, in part, solubility, cohesive energy density, molecular volume, dipole moment, vapor pressure, and dielectric constant.

¹⁵ C. W. Harris, Direct rotation to primary structure, *J. educ. Psychol.*, 39, 1948, 449-468.

¹⁶ F. N. Jones, Olfactory absolute thresholds and their implications for the nature of the receptor process, *J. Psychol.*, 40, 1955, 223-227.

¹⁷ Jones and Jones, *op. cit.*, *J. Psychol.*, 36, 1953, 207-241.

this line of reasoning, indicate one or both of the following: (1) association of site-types on (or in, *cf.* Mullins¹⁸) the receptor-surfaces, or (2) the degree to which the various substances occupy the same sites. It would be possible to argue that part of the common group-variance arises from the convergence of receptor-cells of different types upon the same mitral cell, but at or near threshold a great amount of such overlapping does not seem to be the rule—in the experimental animal at least.¹⁹ It also is interesting to speculate about the possibility that the factors represent the spatial distribution of different receptor-types, but there is no basis on which either to accept or reject such a suggestion.

The results raise certain questions which may be answered by further research. One of these is whether or not substances classified together by factor analysis will interact differently from those that are not. For example, will selective adaptation be greater for substances which are factorially more closely related? Will variables such as eating, which change odor-thresholds, affect factorially different substances differently? Another question is whether or not another analytical methods will yield the same or similar results.

The results reported here obviously leave much to be determined. It may be said that they do show that individual differences in odor-thresholds are systematic, that there are probably many varieties of odor-receptor, and that no extant scheme of odor-classification corresponds to the way in which the human organism classifies odor-thresholds.

Summary. Olfactory recognition thresholds for 20 substances and 84 *O*s were obtained by means of a controlled-blast technique. The thresholds were intercorrelated and the resulting matrix subjected to a centroid factor analysis, followed by mathematical rotation to an oblique solution. The four oblique factors were further analyzed to determine an appreciable general factor. The four group-factors were considered in relation to the chemical and physical properties of the stimuli, with no clear picture emerging. The results imply that (*a*) individual differences in olfactory thresholds are systematic; (*b*) there are probably many varieties of receptor; and (*c*) the results correspond to no existing scheme of odor-classification.

¹⁸ L. J. Mullins, Olfaction, *Ann. N. Y. Acad. Sci.*, 62, 1955, 247-276.

¹⁹ A review of Adrian's work may be found in M. M. Mozell and Carl Pfaffmann, The afferent neural processes in odor perception, *Ann. N.Y. Acad. Med.*, 58, 1954, 96-108. Their conclusion concerning the number of 'primaries' is the same as that in the present paper.

THE GENERALITY OF THE NORMS OF WORD-ASSOCIATIONS

By ALEXANDER M. BUCHWALD, Indiana University

Currently there is a revived interest in word-associations and in the word-association task as part of the general attack on problems of verbal behavior. Several recent studies suggest that word-associations may play an important role in activities such as learning and perception.¹

For both clinical and experimental purposes it has been necessary to collect word-associations from large samples of Ss and to determine for each stimulus-word the cultural frequencies (relative frequencies of occurrence as a response within a sample) of various responses. Several such sets of norms have been gathered and are available.² It is the generality of these norms with which we are concerned here. Knowledge of the generality of norms is necessary in order that we may know whether a given set of norms provides estimates suitable for a particular situation.³

There are, of course, a number of ways in which sets of normative data for word-association may in fact be poor estimates of frequencies for a particular experimental situation. There are such obvious possibilities as sampling error, and the use of experimental Ss drawn from some population other than that for which the norm was established. This paper is concerned with the additional question: Do variations in the techniques used to obtain data lead to variations in cultural frequencies?

The studies reported here deal specifically with two aspects of this question. Experiment I contrasts the frequencies obtained by oral presentation of stimulus-words with those obtained by visual presentation of the same words. Experiment II contrasts frequencies obtained in a multiple-choice situation (the choices being among responses with high frequencies in Experiment I) with those obtained in Experiment I.

* Received for publication October 20, 1955.

¹ J. J. Jenkins and W. A. Russell, Associative clustering during recall, *J. abnorm. soc. Psychol.*, 47, 1952, 818-821; W. M. O'Neil, *The Effect of Verbal Association on Tachistoscopic Recognition*, Technical Report No. 4, ONR Contract N8onr-66216; W. A. Russell and L. H. Storms, Implicit verbal chaining in paired-associate learning, *J. exp. Psychol.*, 49, 1955, 287-293.

² G. H. Kent and A. J. Rosanoff, A study of association in insanity, *Amer. J. Insanity*, 67, 1910, 37-96, 317-390; Russell and Jenkins, *The Complete Minnesota Norms for Responses to 100 Words from the Kent-Rosanoff Word Association Test*, Technical Report No. 11, ONR Contract N8onr-66216, 1954, P. E. Schellberg, A group free association test for college students, Unpublished thesis, University of Minnesota, 1930.

³ Strictly, it is usually not frequencies *per se* that are desired but rather the 'associative strengths,' i.e. the relative probabilities of elicitation of particular words as responses by a stimulus-word, regardless of whether instructions for word-association have been given or not. Associative strengths are, however, commonly taken as proportional to cultural frequencies.

EXPERIMENT I

Subjects. A total of 452 undergraduates (men and women) served as Ss. They comprised the total number of students enrolled in two sections in introductory psychology, present on the day the experiment was conducted, who turned in reports containing one legible response for each stimulus-word. One class ($N = 214$) formed the 'visual' group. The other ($N = 238$) formed the 'auditory' group. The method of assigning students to sections minimized differences between the groups.

Materials. Ten stimulus-words were used that had been selected from a list of 100 words on which associative norms had recently been established.⁴ Three words (*dark, table, boy*) were selected from among words having a frequency greater than 75% for the commonest response. Three words (*bed, eagle, heavy*) were selected from among words having a frequency of approximately 50% for the commonest response and less than 10% for any other response. Two words (*afraid, thirsty*) were selected from among words having frequencies of over 20% for at least two responses. Two words (*comfort, wish*) were selected from among words having small but similar frequencies for at least five responses (range 7 to 12%).

The 10 words were arranged in a list in the following order: *bed, afraid, eagle, comfort, dark, thirsty, table, boy, heavy, wish*. The list, widely spaced, was mimeographed on white typewriter paper. The same order of words was used for all Ss.

Procedure. At the start of the class hour, a blank sheet of yellow paper was given to each S in the auditory group and a copy of the mimeographed word-list, covered by a yellow sheet, was given each S in the visual group with the statement that he was not to look at the words upon it until told to do so.

The following instructions were read to the visual group:

This is part of a word-association experiment. When I tell you to do so, slip the blank sheet down about $1\frac{1}{2}$ in. or until the first word appears. Look at the word. Write down the first word you think of next to the mimeographed word. Each time I say: "Uncover the next word," slide the blank sheet down about $1\frac{1}{2}$ in. and write the first word you think of next to the mimeographed word. Remember, do not look at the next word until I say: "Uncover the next word." Then slide down the blank sheet, read the next word and write down the first word you think of. Are there any questions? Ready. Uncover the first word.

The auditory group was asked to write the numbers I through 10 down the yellow answer sheets. They were then given instructions similar to those for the visual group except that the instructions dealing with uncovering the stimulus-words were omitted and the phrases "say a word" and "hear a word" were used when appropriate. For this group stimulus-words were spoken over a loud-speaker.

A 30-sec. delay occurred between the presentations of the successive stimulus-words. At the end of the session the Ss were requested not to discuss the experiment.

Responses to the various stimulus-words were tabulated for each group in accordance with the following rules: (1) Tabulate responses only for papers with 10 responses. (2) Omit papers where a response cannot be read. (3) If the initial response has been erased or crossed out but can be read, count that instead of the response which replaces it. (4) Each response is counted in a category separate from other responses differing in spelling, except (a) plural nouns are counted with singular forms, and (b) obvious misspellings are counted as correct.

⁴Russell and Jenkins, *op. cit.*, 287-293.

Differences between the frequencies of responses of the two groups were tested for each stimulus-word separately, using $2 \times n \chi^2$ tests. For each stimulus-word an n was used that would yield an expected value of at least 10 in each group, plus one category containing all other responses.

Results and discussion. Frequencies obtained from the two groups differ significantly at the 5% level for 'afraid' and 'eagle' and at the 1% level for 'heavy.' For the other seven stimulus-words there are no significant differences. If these were 10 independent comparisons, these results would reach an over-all significance level of 1%.⁵ Both intuitive considerations and inspection of the data suggest, however, that the assumption that responses to a given stimulus-word are independent of those to other stimulus-words is untenable. Under these circumstances there is no way of testing the over-all significance of differences between groups. Nevertheless, these results suggest the possibility of differences in cultural frequencies dependent on modes of stimulus-presentation.

EXPERIMENT II

Subjects. A total of 432 men and women served as Ss. They comprised two additional introductory sections of psychology. Each class was divided into two subgroups depending upon which of two forms of response-sheet was used. One class provided Ss for the Visual 1 ($N=111$) and Visual 2 ($N=113$) groups. The other class provided Ss for the Auditory 1 ($N=105$) and Auditory 2 ($N=103$) groups.

Materials. The stimulus-words used in Experiment I were used again and were presented in the same order.

For each stimulus-word, except *comfort*, the four responses occurring most frequently in Experiment I were selected to be used as response-alternatives. For *comfort* the five most frequent responses were used since the fourth and fifth were tied. Each alternative was randomly assigned to one of the columns of the response-sheet. The set of alternatives for each stimulus-word was arranged in a row. Two such forms of response-sheet were prepared with the response-alternatives being in a different random order on each. Response-sheets for the visual group had each stimulus-word printed below its appropriate row of responses. In all other details there were no differences between the response-sheets used by the auditory and visual groups.

Procedure. The procedure was similar to that used in Experiment I with the following exception. The Ss were instructed when to expose a particular line of response-alternatives. After 15 sec. they were given the stimulus-word. In addition, instead of ordinary instructions, the Ss were told to look at a set of response-alternatives, and then when they had heard or read the stimulus-word to "draw a circle around the first one of these words you think of."

Responses were tabulated for each stimulus-word and each form separately. Again, the frequencies obtained from different groups were compared by $2 \times n \chi^2$ tests using the same criterion as in Experiment I for the inclusion of a response category in the table.

⁵ Bryan Wilkinson, A statistical consideration in psychological research, *Psychol. Bull.*, 48, 1951, 156-158.

Results and discussion. Chi-square tests were first carried out between frequencies obtained from the two subgroups for each mode of presentation. For the auditory groups all stimulus-words except *table* and *heavy* elicit responses which differ for the two subgroups at the 5% level of significance. For five of these words the differences reach the 1% level of significance. In the visual subgroups, however, for only one stimulus-word, *eagle*, do the responses differ at the 5% level of significance. For the stimulus-words eliciting similar responses on the two forms the frequencies obtained were pooled and compared with the frequencies obtained in Experiment I. Where no significant differences between modes of presentation were obtained in Experiment I, the pooled frequencies for the auditory and visual groups were used. For other stimulus-words the frequencies from the appropriate group were used. All of the 11 Chi-squares are significant at the 1/10% level of confidence except those for *afraid* in the visual group (1%) and *bed* in the visual group (not significant).

The difficulties mentioned in connection with Experiment I concerning the assaying of an over-all level of probability of differences are present here as well. Too, there is presumably some inflation of probabilities due to the use of a preliminary test that excludes some words. In view of the magnitude of differences obtained it would seem safe, however, to assert that differences in frequencies obtained by multiple-choice and common word-association do occur.

Differences between frequencies obtained by the usual word association technique and the multiple-choice technique may have been, in part, due to the fact that these two sets of data were obtained from different, although overlapping, populations. Only those individuals who produce one of these relatively frequent responses were considered in collecting the data. In the multiple-choice technique the entire group of Ss was forced to choose among these same responses.

This explanation cannot, by itself, account for all of the differences. In several instances the frequency of a particular alternative is strikingly smaller, for all groups of Ss, in Experiment II than in Experiment I. For example, whereas approximately 85% of the Ss responded 'girl' to *boy* in Experiment I, less than 50% of the Ss in any subgroup of Experiment II gave this response. If we accept 85% as a reasonable estimate of the 'boy-girl' probability in the word-association situation, no matter which alternative is chosen by the other 15%, the frequency of the 'girl'-response should not decline.

An additional hypothesis to account for the differences between frequencies obtained by the multiple-choice and word-association techniques may be related to what Skinner has referred to as 'prompting' which seems to be a special case of a phenomenon Russell has investigated and has called 'recency'.⁶ Assume that a given S has a hierarchy of responses to each stimulus-word. Presumably frequencies obtained by the word-association techniques depend upon the nature of these individual hierarchies. As Russell and Camp have found, however, if a word has been seen or heard recently, there appears to be an increase in the probability of that word occurring as a response, at least to stimulus-words for which it initially has some

⁶ B. F. Skinner, *Science and Human Behavior*, 1953, 213-214; Russell, Bi-directional effects in word association, SSRC Conference on Word Association, University of Minnesota, 1955.

minimal probability of occurrence.⁷ Accordingly, prior presentation of response alternatives can be assumed to alter the probabilities of occurrence of words in an individual's response hierarchy and consequently to alter cultural frequencies.

A final hypothesis regarding the differences between frequencies obtained by multiple-choice and word-association should be mentioned. There may be a contextual effect induced by the presence of the response-alternatives which serves to increase the probability of some one, or more, alternatives thus altering the relative frequencies of the responses obtained by multiple-choice from those obtained by word-association.⁸

These hypotheses are not mutually exclusive. There is no reason why, in fact, all of the factors hypothesized may not be operating simultaneously.

SUMMARY

In Experiment I, Ss were asked to give a word association to each of a list of 10 words. One group of Ss had the stimulus-words presented visually; the other group had them presented orally.

In Experiment II responses frequently given in Experiment I were presented to Ss in what might be called a multiple-choice association-task. The two means of presenting the stimulus-words and two orders of presentation of the response-alternatives for each stimulus-word were used.

The results suggest the following:

- (1) Visual and auditory means of stimulus-presentation may give different frequencies in word-association.
- (2) Frequencies obtained in a multiple-choice association-task differ with the order in which response-alternatives are written for auditory but not for visual presentation.
- (3) Frequencies obtained from a multiple-choice association differ from those found in a word-association.

⁷ B. W. Camp, The formation and recall of associations as a function of syllable familiarity and frequency of pairing. Unpublished doctoral dissertation, Indiana University, 1954.

⁸ Davis Howes and C. E. Osgood, On the combination of associative probabilities in linguistic contexts, this JOURNAL, 67, 1954, 241-258.

A DESCRIPTIVE ANALYSIS OF THE NON-RANDOMNESS OF SERIAL THRESHOLD-DATA

By WILLARD F. DAY, The Johns Hopkins University

Several experiments have shown that if a stimulus at threshold is repeatedly presented with a short interstimulus-interval, the successive responses are often grouped together in too many long runs of the same response to be a random series.¹ If one is interested only in making the generalization that a series of threshold responses is not likely to be random, it is not necessary to know more about a particular series than that it does not adequately fit a random model. If, however, one is interested in isolating the factors responsible for the lack of randomness, more specific knowledge of what the non-randomness is like must be obtained, *i.e.* some consistent order in the non-random structure of threshold behavior must be identified and described. The research reported here is a descriptive study of the way in which non-randomness actually appears in representative threshold-data. The design of the experiment provides a sample of serial threshold data at a variety of interstimulus-intervals. The stimulus-series consisted of brief flashes of light at the absolute luminance-threshold of vision.

METHOD

Apparatus. The apparatus was a modified Model III Hecht-Shlaer Adaptometer that had previously been used by Verplanck in studies of threshold non-randomness.² A circular stimulus-patch subtending 3° at the cornea was located 7° below a red fixation-point. The binocular patch was illuminated for 0.27 sec. at each stimulus-presentation. The wavelengths of the flash were restricted to a small region near 510 m μ .

O was so housed in a light-tight chamber that when he sat comfortably erect the fixation-point was clearly visible through artificial pupils 2 mm. in diameter. His head was supported and held in position by an adjustable bite-board and brow-rest. At the beginning of each session, O wore dark-adaptation goggles for 25 min. after which he sat in complete darkness for an additional 15 min. He was instructed to keep his eyes on the fixation-point at all times and to respond

* Received for publication February 21, 1956. This report is based on a dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree at the University of Virginia. It was directed by Professor Frank A. Geldard.

¹ W. S. Verplanck, G. H. Collier, and J. W. Cotton, Nonindependence of successive responses in measurements of the visual threshold, *J. exp. Psychol.*, 44, 1952, 273-282; Verplanck, Cotton, and Collier, Previous training as a determinant of response dependency at the threshold, *J. exp. Psychol.*, 46, 1953, 10-14; Collier, Intertrial association at the visual threshold as a function of intertrial interval, *J. exp. Psychol.*, 48, 1954, 330-334; W. F. Day, Serial non-randomness in auditory differential threshold data as a function of interstimulus-interval, this JOURNAL, 1956, 69, 387-395.

² Verplanck, Collier, and Cotton, *op. cit.*, 273-282.

by pressing a key whenever he could tell there had been a flash. No ready-signal was used.

Observers. The *O*s were 11 undergraduates at Harvard University, all inexperienced in threshold-observation. They ranged in age between 17 and 25 yr. and met service visual requirements as tested by the Orthorater.

Procedure. All *O*s were given three sessions of preliminary threshold-training. In the first session, seven intensive levels were determined that bracketed the interval of psychophysical uncertainty. In the second and third sessions, stimuli were presented at these levels by three different psychophysical methods: sensitivity,³ constant stimuli and limits. Changes in method were made without interruption, and a 5-sec. interstimulus-interval was used throughout the 40-min. sessions.

In the subsequent experimental sessions, long series of stimuli were presented at a fixed threshold-intensity, previously determined from the line of best fit to the constants-data from the second and third sessions. Ten of the 11 *O*s served for 6 experimental sessions. In these sessions, interstimulus-intervals of 2, 4, 8, 16, and 32 sec. were varied in counterbalanced order. Two sessions were given at the 4-sec. interval. All sessions lasted for 40 min., yielding series 1200, 600, 300, 150, or 75 stimuli length. For purposes of comparison, *O*₁₁ served for 10 experimental sessions, all with a fixed 4-sec. interstimulus-interval.

RESULTS

Method of analysis. In all, 70 series of threshold response were examined for statistical organization. The *z*-score of the statistics of runs was used as a measure

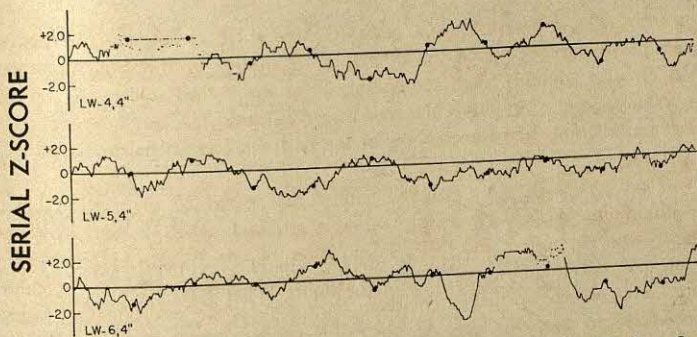


FIG. 1. PLOTS OF SUCCESSIVE SERIAL *z*s FOR SESSIONS 4, 5, AND 6 OF *O*₁₁

of non-randomness, and it was computed with the approximate formulas for the mean and variance of the distribution of number of runs given by Mood.⁴ For binary data, the serial *z* is a function both of the relative frequency of either category and the length of series analyzed. It is not a useful index of non-random organization when there are fewer than 10 members in either category, since the normal

³ W. J. Dixon and F. J. Massey, *Introduction to Statistical Analysis*, 1951, 278.

⁴ A. F. Mood, *Introduction to the Theory of Statistics*, 1950, 393.

distribution is not a tenable approximation to the distribution of number of runs beyond this point. As a result, serial z -scores lose their meaning in evaluating psychophysical data near the borders of the interval of uncertainty.

Plot of successive serial z s. Changes in the non-random structure of the response-series can be described by a graphic plot of successive serial z -scores. These scores were first computed for all possible subseries of 50 consecutive responses in the data, regardless of the length of the response-series for the session as a whole. These z -scores were then serially plotted from left to right in the same order in which the corresponding subseries were derived from the original series. Fig. 1 shows the plot of successive serial z s for Sessions 4, 5, and 6 of O_{11} . Discontinuities in the function correspond to subseries where the number of responses in either

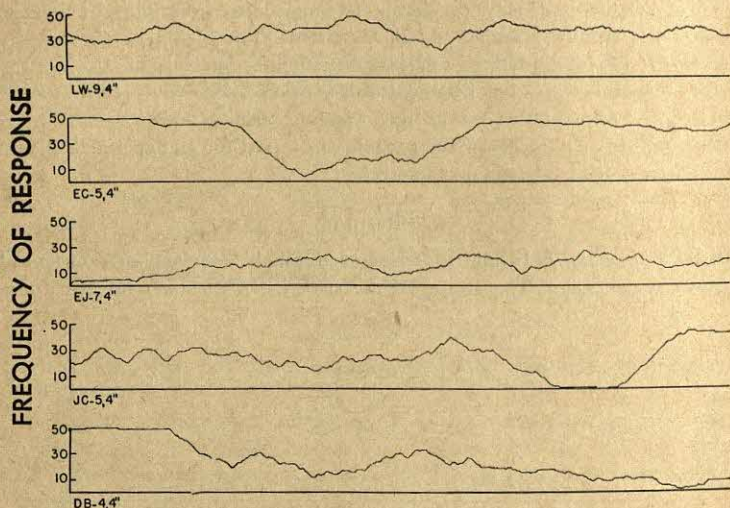


FIG. 2. PLOTS OF SUCCESSIVE FREQUENCIES OF RESPONSE FOR SELECTED SERIES

category was less than 10. The z -scores for adjacent non-overlapping subseries of Length 50 are plotted as large dots in the figure. These series are representative of the data as a whole at all intervals. The series of O_{11} , who observed only at the 4-sec. interstimulus-interval, uniformly show less variability in excursions from the base line at zero. The series for the remainder of the O s were more variable both in the range of z -scores and in their pattern as a function of time.

This kind of descriptive analysis shows that there are by no means always too few runs in series of Length 50. Subseries with too many runs frequently occur. When all the data are considered, cyclical changes in the serial z do not appear, nor is the fluctuation of the z -score restricted to a narrow region of possible values. No trend appears in the serial z as a function of practice, and the plots of successive serial z s differ rather widely in form from day to day and O to O . The only trend to be seen as a function of interstimulus-interval is the somewhat greater range in size of the serial z s, both positive and negative, for the short 2-sec. interval. Taken

as a whole, however, the distribution of serial z s does not fit a random model. Median z -scores for the 20 series at the 4-sec. interval (excluding the data of O_{11}) afford a picture of the overall serial organization of threshold response for a fair sample of comparable data. Since all of these 20 median z s were negative in direction, the hypothesis of randomness can be rejected at the 1%-level by a sign test.

Plot of successive frequencies of response. A similar analysis of the data can be made in search of characteristic changes in relative response-frequency. The frequency of response for all subseries of Length 50 was plotted as a function of serial order. Fig. 2 shows such plots of successive frequencies of response for 5 selected series. Here again no systematic trends appear in the data. For some series, the response-frequency varies only slightly throughout the 40-min. period of observation. For others, the range of response-frequencies covers the entire interval of uncertainty. At times, the response-rate is initially high and becomes progressively smaller as the session continues. At other times, O comes gradually to see more of the flashes as he continues his observation. For most sessions, however, there is no orderly change of rate in either direction.

Serial z s as a function of interstimulus-interval. The use of small sub-series of length 50 is not an efficient way to look for the relationship between interstimulus-interval and serial non-randomness. Table I gives median z s for different interstimulus-intervals where a variety of series-lengths have been used. The z -scores of Row A are based on response to the middlemost 150 stimuli of each session. The median z s for the 2- and 4-sec. interstimulus-intervals are significantly non-random

TABLE I
Z-SCORES FOR SERIES OF VARIOUS LENGTHS
Interstimulus-interval (sec.)

Series-length	2	4(a)	4(b)	8	16	32
(A) 150	-3.234	-3.057	-3.160	-1.582	-1.072	—
(B) 300	-4.836	-5.839	-3.684	-3.547	—	—
(C) 600	-6.885	-8.552	-5.474	—	—	—
(D) 50 (2 subseries)	-0.862	-1.222	-1.493	-1.350	-0.274	-1.350
(E) 50 (all subseries)	-1.33	-1.29	-1.00	-1.38	-0.31	-1.44
(F) 100	-3.015	-2.784	-1.975	-2.867	-0.638	—

in the direction of too few groups. The medians for the 8- and 16-sec. intervals are well within the range of randomness. The difference in size of the serial z for these two groups is significant at the 1% level by the Mann-Whitney ranks test. Such a decrease in the amount of serial non-randomness as a function of interstimulus-interval is directly comparable to the similar results reported by Day and by Collier.⁵ The data of individual O s, however, only very rarely follow the general trend of the median scores, an observation also previously reported by Day.

It is important to note that, with a change in the length of series analyzed, the simple relationship between the amount of non-randomness and interstimulus-interval no longer appears. The z -scores of Row B are based on series of Length 300. The

⁵ Day, *op. cit.*, 389 ff.; Collier, *op. cit.*, 330-334.

medians in Row *C* are for series of Length 600. The medians in Row *D* are based on the first and last subseries of Length 50 in each series. Those in Row *E* are based on the median *z*-score for all subseries of Length 50. Those in Row *F* are for selected 'good' subseries of Length 100, with probabilities of response between 0.45 and 0.55. These results make it difficult to attribute much meaning to the significant difference found at Length 150.

DISCUSSION

The implications of this study are clear with respect to the interpretation of threshold non-randomness. At the present time no meaning can be given to the concept of threshold non-randomness other than to understand simply that as a general rule threshold-response is not apt to fit a random model very well, particularly with short interstimulus-intervals. The inability thus far to identify regularities in the occurrence of non-randomness implies that at present the concept carries no phenomenal reference beyond that of its operational definition. An operational definition of threshold non-randomness can be made in terms of the rejection of specifiable null hypotheses of serial randomness.

Concepts of this kind, that carry no more than operational meaning, are relatively weak starting points for scientific inquiry. Essentially the same amount of meaning is contributed by an arbitrary set of operations. *Converging operations*, of the type described by Garner, Hake, and Eriksen,⁶ convey a relatively more useful meaning. The fact that converging operations do not exist for the concept of threshold non-randomness suggests that it is appropriate to withhold the formulation of explanatory hypotheses until more definitive research has been carried out. In such research, primarily descriptive in nature, the effort should be made to refine an experimental procedure by which certain specifiable statistical characteristics of threshold-response can be counted upon to appear with regularity.

SUMMARY

Seventy threshold response-series, elicited by the repeated presentation of a visual stimulus at the absolute luminance-threshold, were descriptively analyzed for consistency in serial statistical structure. Interstimulus-intervals of 2, 4, 8, 16, and 32 secs. were used. No evidence of consistent serial organization was found. At all intervals the series were at times random, at times non-random in the direction of too few runs, and at other times non-random in the direction of too many runs. The average non-randomness with long interstimulus-intervals was found to be significantly less than that for shorter intervals when series of Length 150 were analyzed. This relationship did not appear when other series lengths were considered. These results were interpreted to mean that at present little behavioral consistency can be shown to underlie the concept of threshold non-randomness.

⁶ W. R. Garner, H. W. Hake, and C. W. Eriksen, Operationism and the concept of perception. *Psychol. Rev.*, 63, 1956, 149-159.

CHARACTERISTICS OF MENTAL IMPAIRMENT IN HYPOXIA

By CLINTON H. MAAG, Office of Naval Research

Previous investigations of the effect of hypoxia on performance, competently reviewed elsewhere, have suggested that efficiency declines gradually and continuously with duration of exposure to an oxygen-deficient environment until loss of consciousness ensues.¹ An alternative view is suggested by the physiological finding that a stepwise decrement in oxygen saturation is characteristic of adjustment to oxygen deficiency. Following an initial decline, homeostatic mechanisms come into play (increased blood pressure, ventilation, ventricular dilation, output of hemoglobin, and the like) which maintain a constant supply of oxygen to the brain, until finally, at the compensatory limit, these mechanisms become inadequate to maintain homeostasis and unconsciousness supervenes. Why should there be a continuous decline in the efficiency of performance during a period of physiological constancy? The hypothesis to be investigated here is that there is a steady state of efficiency which follows an initial decline and persists until the time of collapse. This steady state is characterized by uniform performance interspersed with brief periods of inadequate performance called *blocks*.² These periods of impaired performance or blocks increase in frequency with the duration of stress, giving the appearance of a continuous decline in efficiency when level of performance is averaged over relatively long temporal intervals. In the present study, measures of performance were corrected for blocking.

Subjects. Volunteers were requested from the enlisted personnel of the Naval Air Basic Training Command, and 10 men were selected from this group to participate in the experiment. All were within the age-range of 18-20 yr., in good physical health, of average intelligence, and without apparent personality disorder.

Task. Most studies of behavior under conditions of hypoxia have been concerned with sensory processes, psychomotor skills, and simple mental tasks. The present study sought to evaluate the efficiency of higher level mental processes in an hypoxic environment. The measuring instrument utilized, similar to the concept-formation task of Hanfmann and Kasanin,³ is referred to as the Conceptual Reasoning Test (CRT).⁴ In brief, the CRT consists of 32 small wooden blocks which

* Received for publication March 7, 1956. This paper is adapted from a doctoral dissertation presented to the University of Florida in 1954.

¹ Tufts College Institute of Applied Experimental Psychology, *Handbook of Human Engineering Data for Design Engineers*, 2nd ed., Part VII, 1951, 1-15; R. A. McFarland, The psychological effects of oxygen deprivation (anoxemia) on human behavior, *Archiv. Psychol.*, 22, 1932, 1-135; H. G. Armstrong, *Principles and Practice of Aviation Medicine*, 3rd ed., 1952, 210.

² Cf. A. G. Bills, Blocking: A new principle of mental fatigue, this JOURNAL, 43, 1931, 230-245.

³ Eugenia Hanfmann and Jacob Kasanin, A method for the study of concept formation, *J. Psychol.*, 3, 1937, 521-540.

⁴ C. H. Maag, Development and evaluation of a conceptual reasoning test, *Educ. psychol. Measmt.*, in press.

embody the five dichotomous characteristics of: large-small, tall-short, round-square, hollow-solid, and numbered-lettered. Half, *i.e.* 16, of the blocks are similar in at least one characteristic. For instance, half are numbered, half are lettered, and so forth. Eight of the blocks are similar in at least two characteristics; *e.g.* tall-hollow, square-short, or numbered-large. Four of the blocks are similar in at least three characteristics; *e.g.* short-solid-numbered. Two of the blocks are similar in four characteristics; *e.g.* large-short-numbered-hollow. No two blocks are similar in all five characteristics. Utilizing any one or all five dichotomous characteristics of the blocks, one may classify the blocks in any of several different ways. It is the task of *S* to determine the classificatory system utilized by *E*. The lowest level of difficulty would be a system utilizing a single characteristic such as tall, round, or hollow. A more difficult series of problems are those involving two-category classifications, such as tall-hollow or round-numbered.

On the basis of preliminary experimentation with this test, it was decided to use those problems involving classification by means of only two characteristics. There are 40 dual-classification problems that may be set for the 32 blocks. *E* had a list of these 40 possible classifications with space available for recording *S*'s order of selection, time, and errors. It will be recalled that any 8 of the 32 blocks may be classified by two of the characteristics, and *E* may use any one of them as a model which he places before *S*. It is the task of *S* to determine, by consecutive testing of hypotheses, the remaining seven blocks having the pair of characteristics in common with the model. Since the model embodies five characteristics, the only information given is that the correct solution must involve a pair of the five characteristics. Once *S* has learned the proper procedure for establishing and testing hypotheses to determine the correct pair, he reaches a relatively constant level of performance in terms of time and errors.

Procedure. To prevent the effects of hypoxia from being comouflaged by an improvement in performance as a result of further practice, a preexperimental training period was employed. The period of training continued until all *S*s reached a plateau. The final 80 trials of the preexperimental training series were grouped into four sections of 20 consecutive problems for each *S*. An analysis of variance was made on these data which indicated a constant level of performance had been attained. The procedure for determining reliability proposed by Hoyt,⁵ and later amplified by Hoyt and Stunkard,⁶ provided a reliability coefficient of 0.9795.

Since periods of inadequate or impaired performance under stress are postulated, it is necessary to establish a criterion of a typical response or blocking. It is reasonable to suppose that performance under stress will deviate from *S*'s norm in a positive direction; *i.e.* he will require a longer time to solve the problems. Since we are interested in only one-half the normal curve, we may designate as a block any response which exceeds *S*'s plateau-mean by 3.09 *SD*. Although this is a somewhat arbitrary cut-off point, it puts a rather heavy burden on the hypothesis to be tested, since it is quite possible that there could be periodic blocks within the limits imposed. If so, mean performance would increase with duration of stress,

⁵ C. J. Hoyt, Test reliability estimated by analysis of variance, *Psychometrika*, 6, 1941, 153-160.

⁶ C. J. Hoyt and S. L. Stunkard, Estimation of test reliability for unrestricted item scoring methods, *Educ. psychol. Measmt.*, 12, 1952, 574-587.

and one would be forced to reject the blocking hypothesis. On the other hand, the cut-off point must be sufficiently liberal to allow significant differences in performance to be demonstrated if such differences are present.

S was seated opposite *E* in a double-lock decompression chamber. *E* wore a demand-type oxygen mask throughout the experimental session. A similar mask was available to *S* when he became unconscious or an emergency arose. A physician was available for immediate service at all times during the periods at altitude. There were five experimental sessions in the decompression chamber: sea level, 13,000, 16,000, 17,000 and 18,000 ft.

S was first exposed to a simulated altitude of 12,000 ft. in a session primarily designed to acquaint him with the procedures, to establish a smooth mode of operation, and to relieve any general anxiety that might influence performance in the actual experimental sessions. A sea-level session served as a control for variables extraneous to degree and duration of hypoxia. Actually, the sea-level session was at an altitude of approximately 150 ft. To keep the sea-level conditions as comparable as possible, the chamber was initially brought up to 10,000 ft. and immediately dropped to 150 ft. (simulated altitude). Since with this procedure it was necessary to keep all pumps and ventilating equipment in operation, noise and other variables were comparable with conditions at other altitudes.

The initial period of ascending to altitude was used by *S* as a 'warming up' period for the conceptual reasoning test (*CRT*). As soon as the simulated altitude was reached, *S*'s performance was recorded. *S* continued to work until collapse, at which time he was immediately revived with oxygen and no longer took part in the experiment for that particular session.

Since all our *Ss* did not become unconscious at the same time, it was impossible to compare performance over temporally equal working periods. For instance, the more stable *Ss* continued to perform for more than an hour while the less stable collapsed during the first half hour at altitude. Thus a comparison of performance-decrement over successive periods of time would have very little meaning. For this reason a procedure was used to make the performance of the different individuals comparable at the same relative points in their decline. All the responses of each *S* were divided into five phases, each phase containing the same number of responses. The method is similar to that devised by Vincent for combining individual learning curves.¹

Although it might seem appropriate to test the hypothesis of blocking with reference to degree as well as duration of hypoxia, the design of the present experiment would make interpretation rather difficult. The testing sessions were conducted over a period of 14 days which, because of high day-to-day variability in susceptibility to hypoxia, would contaminate data comparing performance at different altitudes. Owing to this factor and other variables of procedure, the present paper will be restricted to an evaluation of performance from normalcy to unconsciousness under constant degrees of stress.

Results. The data obtained from this experiment are summarized in Table I. It may be noted that the *F*-ratios for the session at sea level indicate chance differences among the five phases of performance. This situation facilitated analysis con-

¹ S. B. Vincent, The function of the vibrissae in the behavior of the white rat, *Behav. Monogr.*, 1, 1912, 28-32.

siderably since it did not necessitate an elaborate statistical treatment of the data to filter out the effects of fatigue, boredom, motivation, and other contaminating factors.

The uncorrected data (including blocks) indicate a progressive decline in performance at all altitudes. The differences in mean performance among the five phases are significant at better than the 1% level of confidence for the three highest simulated altitudes and the 5% level for 13,000 ft. These findings fit the commonly accepted generalization that performance gradually deteriorates with time in an hypoxic environment.

One might question the utility of correcting the data for blocking. It is possible that chance differences would exist if one merely eliminated the first phase from consideration. The answer to this question is provided by a separate analysis of the uncorrected data for the last four phases of each session (F_{2-5}). Table I indi-

TABLE I
SUMMARY OF DATA BY ALTITUDE AND PHASE

		Phase					F_{1-5}	F_{2-5}
		1	2	3	4	5		
Sea level	Number of responses	295	293	292	291	287		
	Mean time per response	17.48	17.05	17.23	16.98	16.84	2.39	5.09
	Percentage of blocks	6.10	3.75	3.42	4.47	2.79		
	Mean time excluding blocks	16.35	16.33	16.74	16.10	16.29	1.75	1.15
13,000 ft.	Number of responses	277	277	272	270	267		
	Mean time per response	16.77	17.64	17.75	18.39	18.37	2.47*	1.20
	Percentage of blocks	2.61	3.21	4.04	7.04	8.61		
	Mean time excluding blocks	16.09	16.82	17.05	16.92	16.74	1.36	5.93
16,000 ft.	Number of responses	173	173	169	169	166		
	Mean time per response	17.46	18.43	18.49	19.56	21.42	7.13†	6.02†
	Percentage of blocks	4.05	4.62	7.10	12.43	18.07		
	Mean time excluding blocks	16.52	17.68	17.16	17.62	18.49	3.04*	1.67
17,000 ft.	Number of responses	150	147	144	144	143		
	Mean time per response	17.61	18.53	20.65	20.85	21.64	4.44†	2.69*
	Percentage of blocks	6.00	6.80	14.58	13.89	19.58		
	Mean time excluding blocks	16.09	17.15	17.29	17.51	17.83	2.34*	2.21
18,000 ft.	Number of responses	170	166	165	162	161		
	Mean time per response	17.83	19.48	20.24	21.27	23.00	6.41†	3.51†
	Percentage of blocks	9.41	7.83	9.09	12.96	18.01		
	Mean time excluding blocks	16.24	17.90	18.46	18.19	19.04	5.38†	1.06

* Significant at the 5% level. † Significant at the 1% level.

cates that significant differences exist among these phases for the three highest altitudes. At 13,000 ft., however, omission of the first phase shows chance variations in performance. Thus, even at this point in the analysis one might question the generality of the concept of progressive decline, at least for milder degrees of hypoxic stress.

It was hypothesized that progressive deterioration is a consequence of deviant responses or blocks which increase in frequency with duration of exposure. Table I shows the incidence of blocks for each phase of performance. It may be noted that, with the exception of the sea-level session during which the frequency of

blocks decreased, all periods at altitude show an increase in the percentage of blocks. Although there are reversals at several points, it is clear that the trend of the data is for frequency of impaired responses to increase with duration of stress. As one might expect, a tabulation of the error-scores shows a corresponding increase with duration of stress.

We may now determine the effect of eliminating the blocks from the data. It is recognized that without a criterion for what shall be considered as impaired response, the mean performance could be made to increase, decrease, or even remain constant as a function of the amount of data excluded. It will be recalled that only those responses which exceeded the individual's plateau-mean by 3.09 SD were considered as deviant. Consequently, the means are free to increase up to the criterion-limits, which provides a sufficient range to test for chance-fluctuations of performance. An examination of the *F*-values for all five phases of the corrected data (blocks omitted) indicates chance-differences for the 13,000-ft. session. The performance for the other altitudes, however, shows a significant difference among phases at better than the 5% level of confidence. Omission of blocks alone does not, therefore, indicate a constant level of performance. One also must take into consideration the rapid decline during Phase I while *S* is becoming hypoxic prior to the steady state achieved by the compensatory mechanisms. The results of this analysis are presented in Table I in terms of *F*-values for the corrected data of the last four phases. Only chance differences exist for all altitudes. In the light of this analysis, we cannot reject the null hypothesis, and the concept of progressive decline therefore seems a less accurate description of behavior under hypoxia than does the concept of initial decline and intermittent impairment. It should be noted that it takes relatively few deviant responses to produce the impression of gradual deterioration in performance when no correction is made for them. For instance, during the final phase of performance just preceding collapse, approximately 80% of the responses at the two highest altitudes were within the criterion-limits. Correction of the data to take account of these blocks reconciles performance with the concept of homeostasis and physiological constancy.

Discussion. A description of behavior analogous to that observed in the present experiment is provided by Broadbent, who asserts that "when a man has been working for some time his performance becomes irregular rather than breaking down completely. Crudely speaking, a man is not like a child's mechanical toy which goes slower as it runs down. Nor is he like a car engine which continues normally until its fuel is exhausted and then stops dead. He is like a motor which after much use misfires, runs normally for awhile, then falters again, and so on."⁴ The generality of the findings offered in this paper is open to experimentation using other tasks and other types of stressful situation, such as extreme temperature, noise, and loss of sleep.

⁴D. E. Broadbent, Variations in performance arising from continuous work, *Conference on Individual Efficiency in Industry*, Medical Research Council, 1955, 1-5.

MEMORY FOR VISUAL FIGURES BY THE METHOD OF IDENTICAL STIMULI

By LAWRENCE KARLIN and GLEN BRENNAN, New York University

This investigation was prompted by two recent studies of memory for visual form by the method of identical stimuli. Crumbaugh, using a variety of figures and time-intervals ranging from 0-12 sec., found progressive effects which he interpreted as supporting the Gestalt concept of autonomous change.¹ Seminara, in an earlier, unpublished study, found no significant time-errors for ellipses of two sizes with an interval of 8 sec.² In the present investigation, time-errors are determined for a broken circle and for an ellipse.

The principle of closure suggests an increasingly negative time-error for the circle, but the prediction for the ellipse depends on which dimension is judged. If the trace of the ellipse becomes circular in time, judgments of width should give an increasingly positive time-error, while judgments of length should give an increasingly negative time-error. Judgments of width are used here because, while artifacts of method might be expected to produce similar time-error functions for both the circle and the ellipse, they could hardly be expected to produce the dissimilar functions predicted. Of course, there are other possible combinations of change in the dimensions of the trace of the ellipse which could result from decreasing eccentricity, but it seems desirable to begin with the simplest possibilities.

METHOD

Apparatus. A 500-w., 14-in. Bausch and Lomb opaque projector was used to project the stimulus-figures, and a 300-w., 5-in. TDC, 2 × 2 slide-projector maintained the level of illumination during rest-periods and interpolated intervals. A solenoid-operated, double-leaf tachistoscope was so mounted that as one leaf uncovered one projector, it simultaneously covered the other. All time-intervals were controlled by four Hunter timers arranged in series.³ This apparatus was contained in one room, and the stimulus-figures were projected through a glass window on a beaded screen, 40 in. wide and 30 in. high, located in the adjoining room. A telegraph-sounder in this room provided a ready-signal 2 sec. before every projection.

Stimulus-figures and conditions of presentation. The figures for the main experiment were black line-drawings on white, matte board. The projection of the broken circle had an inside diameter of 7.7 in. and a gap of 15°. The gap was centered on

* Received for publication June 9, 1956.

¹ J. C. Crumbaugh, Temporal changes in the memory of visually perceived form, this JOURNAL, 67, 1954, 647-658.

² J. L. Seminara, An investigation of the time-errors for ellipses and circles relating to the Gestalt theory of memory changes, M.A. thesis, New York University, 1952.

³ Lawrence Karlin, An improved circuit for connecting the Hunter timer and similar timers in sequence, this JOURNAL, 67, 1954, 358-359.

the horizontal meridian and always faced the fixation-point (1 in. in diameter), which was in the center of the screen.

The projected ellipse had a vertically oriented major inside diameter of 9.2 in., and a minor inside diameter of 3.1 in. The centers of each projected figure was located either to the right or left of, and 8 in. away from, the fixation-point. The line-thickness of each figure was 0.6 in. S was seated 11.5 ft. from the screen with his head in a headrest that kept his eye-level about 3.5 ft. above the floor, which was about the height of the center of the screen. The room was in darkness except for the screen, which was illuminated at all times. The rest-period and interpolated interval-brightness measured with a Macbeth illuminometer was 0.039 ft.-L, while the brightness of the stimulus-field was 0.027 ft.-L.

Procedure. Each of 25 experimentally naïve college students served for one session. Fifteen judged the circle and 10 judged the ellipse. The same figure was used both as the standard and comparison, with no objective difference ever being involved. S was told to judge whether the width of the gap in the second of each pair of successively presented circles was greater than, equal to, or less than the width of the gap in the first circle, or similarly instructed with respect to the width of the second of each pair of ellipses. The figures were presented to the right of the fixation-point for half of the Ss and to the left for the remaining half. The instructions were illustrated by a card containing sample figures. The pairs of identical figures were then presented in blocks of 10 trials with interpolated intervals between successive presentations of 1, 4, and 8 sec. These intervals were given in counterbalanced order of 1 4 8 8 4 1 for a total of 20 trials per interval and 60 trials per session, which lasted about 20 min. The rest-period was constant at 15 sec. During the entire session, S's head was in the headrest. He was instructed to keep his eyes on the screen at all times and to look steadily at the fixation-point from the time he heard the warning signal, 2 sec. before the end of each rest-period, until he gave his judgment.

RESULTS

Table I summarizes the results for both figures, Lauenstein's $D\%$, as defined in Table I, was used to measure the time-error.⁴ It measures the preponderance of less judgments relative to the total number of judgments, and its sign is negative with an excess of *greater* judgments.

Circle. The combined values for the broken circle become increasingly positive as the time-interval increases. Using a distribution-free method described by McNemar for sets of correlated measures,⁵ a Chi-square of 3.45 with two *df.* was obtained, which is not significant at the 5% level, but this method may not provide an adequate test, because four individuals could not be used and insignificant differences in one region may have obscured significant differences in another. This reasoning is supported by the results of the sign-tests obtained for all three pairs of interpolated intervals. For the comparison of the 1- and 4-sec. intervals, if the two zero differences are ignored, 11 out of 13 differences were in the same direction. The chance-probability of this result is about 1%. The comparison of the 1- and 8-sec. intervals yielded 13 out of 15 differences in the same direction. The

⁴ Kurt Koffka, *Principles of Gestalt Psychology*, 1935, 469.

⁵ Quinn McNemar, *Psychological Statistics*, 2nd ed., 1955, 358-359.

chance-probability of this result is less than 0.4%. A like comparison of the 4- and 8-sec. intervals shows no significance. Here, only 9 of 13 differences were in the same direction, and the chance probability of this result is about 9%. The sign-test also suggests that the time-error at each interval considered separately differs significantly from zero ($p < 4\%$ at the 1- and 4-sec. intervals, and $< 1\%$ at the 8-sec. interval).

Since the ascending series always preceded the descending series, it is not possible to separate the effects of repetition from the effects of the particular sequence used. The joint effects of these two factors may be evaluated by comparing the average time-error at each interval for the first half (ascending series) and the second half (descending series) of each session. It is apparent from inspection of this table that the trend in both halves, as well as the signs of the time-errors, remain the same. It should be noted, however, that there is an

TABLE I

AVERAGE TIME-ERROR (IN $D\%$) FOR EVERY INTERPOLATED INTERVAL
 $(D\% = 100 (L-G)/L+G+E$, where L equals the number of judgments of 'less'; G , of 'greater'; and E , of 'equal'.)

Series	Broken circle			Ellipse		
	1 sec.	4 sec.	8 sec.	1 sec.	4 sec.	8 sec.
Ascending (first half)	-3.3	9.3	21.3	-26	-11	-7
Descending (second half)	-5.3	8.0	18.7	0	4	-6
Combined	-4.3	8.7	20.0	-13.0	-3.5	-6.5

algebraic decrease in the time-error from the first to the second half, for each interpolated interval. Despite this consistency, sign-tests for the decrease are not significant.

Ellipse. The results for the ellipse display no consistent trend. Although all three combined time errors are negative, the sign-test suggests that none of the three is significantly so. Inspection of the table also shows no agreement between the results for each half of the session, although the time-errors are consistently less negative or more positive at all three intervals in the second half. The effect on the time-error of factors associated with the first and second halves thus seems to be opposite to the slight effect found for the circle, in that the ellipse yielded an algebraic increase in the time-error rather than a decrease. It cannot be concluded from the present data, however, that these effects are systematic, since sign-tests comparing the first and second halves of each session for both ellipse and circle are not significant.

DISCUSSION

The data of Table I refute the hypothesis that closure is a factor in determining the fate of the trace of the gap in the broken circle, since these results can be interpreted to mean that the gap of the trace is steadily increasing in size with time. The insignificantly negative time-errors for the ellipse, coupled with the lack of any significant trend, also provide no support for the Gestalt thesis. Any trend to decrease in eccentricity would be reflected by an increase in trace width, which in turn would be revealed by increasingly positive or decreasingly negative time-errors.

Comparing our results with those of Crumbaugh shows both similarities and differences. He obtained time-error functions for two gapped circles, one with a gap of 15° , the other with a gap of 90° . For the smaller gap, the results were similar to ours in showing a positive time-error at all intervals. This function was however, unlike ours in showing a reversal in the initial tendency to become increasingly positive at the longest interpolated interval of 12 sec. Like those of the present experiment, these results do not support Gestalt theory, since all time-errors were positive. Furthermore, the reversal in trend which he obtained does not support the Gestalt view that the changes in the trace are progressive.

Crumbaugh's results for the circle with a 90° gap show, however, a progressive trend toward increasing negativity, with all time-errors negative. These results support the Gestalt hypothesis but differ from the results of the present study and also from the results Crumbaugh obtained for the smaller gap. It is difficult to say why an increase in the size of the gap should have yielded such a marked difference in his results. Perhaps the 90° gap so changed the appearance of the figure that it was no longer perceived as an incomplete circle but as some other figure—as, for example, the letter C or G and their mirror-images.

The marked positively directed trend obtained for the broken circle in the present experiment suggests the operation of the principle of emphasis or 'sharpening' noted in previous studies.⁶ As the term implies, this principle refers to the tendency to exaggerate any unusual or distinctive aspect of a remembered situation. The gap may be such a distinctive feature because the instructions emphasized it and because it was the only irregularity of an otherwise regular figure. It would be of interest to determine whether similar time-error functions can be found, using the present procedure, for other regular figures (e.g. triangles, squares) which possess a gap or an indentation.

Seminara's results for the ellipse agree with ours in showing no significant tendencies. This was true both for judgments of length as well as of width and for two ellipses of different eccentricities. It may be that the present results for the ellipse are not significant as in the case of the broken circle because the ellipse is a 'good' symmetrical figure as it stands and would, therefore, not be internally stressed to change in the direction of decreasing eccentricity. In addition, the absence of any irregularity or distinctive feature minimizes the possibility of the sharpening or emphasis that may have operated in the broken circle, which could account for the absence of a significant trend such as was found for the broken circle.

Crumbaugh's results for the ellipse are difficult to evaluate. Inspection of his time-error function does not disclose a clear tendency for the time-error to change for the first four intervals of 0, 0.3, 2, and 5 sec. Only for the last interval of 12 sec. is the time-error significantly negative. Here the negative time-error implies decreasing eccentricity since Crumbaugh's Ss judged the length of the major axis. It is not certain that his results display a significant trend, but if they do, the difference between his and the present results may be due to any of a number of

⁶Frederich Wulf, Tendencies in figural variations, in W. D. Ellis, *Source Book of Gestalt Psychology*, 1950, 136-148; E. V. Fehrer, An investigation of the learning of visually perceived forms, this JOURNAL, 47, 1939, 187-221; J. J. Gibson, The reproduction of visually perceived forms, *J. exp. Psychol.*, 12, 1929, 1-39.

methodological differences between the two experiments, which include: the direction of figure-ground brightness (Crumbaugh used luminous figures on a dark ground); the proximity of field contours to the figures; the shape of the stimulus-field; use of an equality category versus forced judgment of differences; the presentation of standard and comparison figures on same versus opposite sides of the fixation-point; method of exposure (the use of a shutter versus direct switching of the lamp); and randomized versus non-randomized sequence of time-intervals.

It is well to emphasize that the differences between the results of the present method and those of earlier experiments suggest the possibility that the trend found in the present experiment may be due to methodological factors rather than to some characteristic of the memorial process. In fact, the entire history of this problem indicates that adequate methods of studying memory for form have not been worked out. Thus the present method has the disadvantages that accrue from repetitive presentation of stimuli. Such repetition could result in (1) a consolidation or distortion of the memory-trace which would interfere with its temporal course, and (2) a change in the sensory and corresponding cortical surfaces which might systematically alter perception.

In this connection, it should be noted that the trends for the first and second halves of the experiment remain unchanged for the broken circle. In a similar comparison, no significant changes were found for the ellipse, although the consistency was not so clear as those obtained for the circle. These results suggest that repetitive effects may not be important in the sense just mentioned, although this conclusion is limited by the previously mentioned confounding of ascending and descending series with the first and second halves of the session. In any case, to minimize these (hypothetical) defects, it may be necessary in future experiments to decrease the number of repetitions and increase the number of Ss, as was previously done in connection with other methods.⁷ Since such a change magnifies the effects of individual idiosyncracies, it can only complement the present method, not replace it.

Finally, the possibility remains that the much shorter interpolated intervals used both by Crumbaugh and in the present study may be able to reveal properties of the memorial process which do not obtain over the much longer intervals used in earlier studies.

SUMMARY

With the method of identical stimuli and a broken circle and an ellipse, as stimulus-figures, a significant trend in time-error was obtained for the former and insignificant results for the latter. It was suggested that the differences between these findings and those of a previous experiment may have been due to differences in the methods of presenting the figures. In general, the results of the present study did not confirm the Gestalt hypothesis concerning the operation of autonomous factors in the memory-trace. The possibility that the principle of emphasis or 'sharpening' was a factor in determining such changes as were found was considered.

⁷ R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, rev. ed., 1954, 773-776; C. E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 588-593; N. G. Hanawalt, The method of comparison applied to the problem of memory change, *J. exp. Psychol.*, 43, 1952, 37-42; D. O. Hebb and E. N. Foord, Errors of visual recognition and the nature of the trace, *ibid.*, 35, 1945, 335-348.

INCIDENTAL AND INTENTIONAL MEMORY FOR LIFTED-WEIGHTS

By IRVING J. SALTZMAN, Indiana University

Under most conditions, intentional learning is found to be superior to incidental learning. Ordinarily, intentional learners do something which incidental learners do not do, although just what they do that produces the superiority never has been established. It is clear, however, that the process in question takes time; the superiority of intentional learning depends on the rate of presentation of the learning material, on the number of learning trials, and on the orienting task.¹ One of the things that intentional learners might do that incidental learners most likely do not do is rehearse the material, both during and between trials. It is interesting to note that almost all of the factors shown to reduce or to eliminate the superiority of intentional over incidental learning might very easily have interfered with efforts to rehearse.

If the superiority of intentional learning is primarily a function of opportunity for rehearsal by the intentional learners, then the use of a learning task which makes rehearsal difficult should show no superiority of intentional learning under conditions which, with other tasks, ordinarily show such superiority—i.e. a slow presentation rate and several learning trials. Such a task was used in the studies to be described here. Groups of incidental and intentional learners were given varying numbers of trials or opportunities to lift and replace two differently weighted 'standard' cylinders. Following these learning trials, the *O*s were asked to judge other weights in relation to the two standards. The intentional learners were instructed, before they lifted the standards, to try to remember how they felt, how much they weighed; the incidental learners, of course, lifted the standards without these instructions. On the assumption that the intentional learners would find it impossible, or at least very difficult, to rehearse in this situation, it was expected that they would do no better than would the incidental learners.

Experiment I. The *O*s in Experiment I were 120 undergraduate students randomly divided into two groups of 60 each—an incidental and an intentional learning group.² Each group was further subdivided into 3 subgroups of 20 *O*s each, and each subgroup received a different number of opportunities to lift each of two standard weights either one, three, or five times.

* Received for publication June 27, 1956.

¹ I. J. Saltzman, The orienting task in incidental and intentional learning, this JOURNAL, 66, 1953, 593-597; Edith Neimark and I. J. Saltzman, Comparisons of incidental and intentional learning under different rates of stimulus presentation, this JOURNAL, 66, 1953, 618-621; I. J. Saltzman and Rita L. Atkinson, Comparisons of incidental and intentional learning after different numbers of stimulus presentations, this JOURNAL, 67, 1954, 521-524; Saltzman, Comparisons of incidental and intentional learning with different orienting tasks, this JOURNAL, 69, 1956, 274-277.

² The number of *O*s used was 133, but 13 were eliminated for failure to satisfy a criterion which will be described below.

The apparatus consisted of a classroom chair, *i.e.* a chair with an attached arm or table for writing, in which *O* sat, a chair and table for *E*, and nine identical wooden, weighted cylinders. Each cylinder was 86 mm. high and 41 mm. in diameter. In the cylinders were varying amounts of buckshot imbedded in wax. The cylinders weighed 81, 119, 131, 144, 158, 174, 191, 210, and 307 gm. The lightest cylinder and the heaviest cylinder were used as standard weights and the remaining seven were used as test-weights. Four circles, 45 mm. in diameter, were drawn in white chalk on the arm attached to *O*'s chair. For expository purposes these circles will be referred to here by number, although they were never referred to by number in the experimental situation. Circle 1 was located in the upper left-hand corner of *O*'s table, Circle 2 in the lower left-hand corner, Circle 3 in the center of the table, and Circle 4 in the upper-right hand corner. Circles 3 and 4 were empty, but within Circles 1 and 2 were pasted small, differently colored, paper figures. A blue dot, a white triangle, and a yellow square were in Circle 1 and a white dot, a yellow triangle, and a blue square were in Circle 2. At the start of the experiment the two standard weights (81 and 307 gm.) were placed in Circles 3 and 4, respectively. *E* sat in his chair beside *O* but facing in the opposite direction. On a table immediately in front of *E* and out of view of *O* were located the seven test-weights and the data-sheets.

The *O*s in the incidental learning group who received only one learning trial, as soon as they were seated, were told that *E* wished them to follow some simple instructions before the experiment was begun which would familiarize them with the procedure. They were told that the experiment would involve the use of the cylinders that were on their table (in Circles 3 and 4) and that *E* would refer to the colored figures in the other circles (Circles 1 and 2) drawn on the table. The weights were never referred to as weights, only as cylinders. *O* was instructed to keep his right elbow on the table and with the finger tips of his right hand to lift "this cylinder" (*E* pointed to the light weight in Circle 3) about an inch or two from the table, to hold it briefly, and then to place it in the "circle with the blue dot [Circle 1] and leave it there." *E* then moved the other standard cylinder, the heavy one, from Circle 4 to Circle 3 and asked *O* to lift it, hold it briefly, and then place it in the "circle with the white dot [Circle 2] and leave it there." For the *O*s in the incidental learning group who received three or five trials, the procedure was the same with the exception that on the first two trials for the three-trial group and on the first four trials for the five-trial group, as soon as the light weight was placed in Circle 1, it was removed by *E* and held until the next trial was begun; and as soon as the heavy weight was placed in Circle 2, it was removed by *E* and replaced in Circle 4. Each trial, then, began the same way, with the light weight in Circle 3 and the heavy weight in Circle 4. *O* always was so instructed that he placed the light weight in Circle 1 and the heavy weight in Circle 2. On the first trial for all subgroups, Circle 1 was referred to as the circle with the blue dot and Circle 2 as the circle with the white dot. On the second trial for all *O*s receiving more than one trial, Circle 1 was referred to as the circle with the white triangle and Circle 2 as the circle with the yellow triangle. On the third trial for the three-trial and five-trial subgroups, Circle 1 was called the circle with the yellow square, and Circle 2 the circle with the blue square. For the *O*s receiving five trials, Circle 1 was referred to as the circle with the blue dot on Trial 4 and as the

circle with the white triangle on Trial 5, Circle 2 was called the circle with the white dot on Trial 4 and the circle with the yellow triangle on Trial 5.

After the learning trials were completed, *Os* in the incidental learning group were informed that the rest of the experiment had to do with the weights of the two standard cylinders that they had lifted. They were told that shortly they would be asked to lift seven other similar test-cylinders, one a time, and to judge whether each was lighter than the light standard, in which case they were to say 'light,' heavier than the heavy standard, in which case they were to say 'heavy,' or intermediate in weight between the two standards, in which case they were to say 'in between.' *Os* were not told how many of the test-weights would fall into each category.

The seven test-weights were then presented, one at a time, each weight was judged, and *Os*' judgments were recorded. The same random order of presentation of the seven weights was used for all *Ss*. The order was as follows: 191, 144, 210, 174, 131, 119, and 158 gm. The light standard weighed 81 gm. and the heavy one weighed 307 gm. All of the test-weights belonged to the category 'in between,' since none was lighter than the light standard nor heavier than the heavy standard. *E* always placed the test-weights in Circle 3. *O* lifted the weight, made his judgment and replaced the weight in Circle 3. As soon as *O* replaced the weight, *E* removed it and replaced it with the next weight. The time between successive judgments was approximately 5 sec.

After all seven weights had been lifted and judged, *E* moved the light standard from Circle 1 to Circle 3 and told *O* to lift it once more and replace it. *E* then returned the light weight to Circle 1 and moved the heavy standard from Circle 2 to Circle 3 and told *O* to lift it once more. Then the heavy weight was replaced in Circle 2, and *O* was told that *E* wished him to make his judgments of the seven test-weights again. The seven weights were presented a second time in exactly the same manner and in exactly the same order and *O*'s judgments were again recorded. This second series of judgments was included to provide a means for eliminating deviant *Os* who were incapable of a reasonable degree of accuracy in judging the weights. Following this second series, the *Os* were dismissed.

The procedure used with *Os* in the intentional learning group was identical with that for the incidental *Os*, except that the intentional learners were told before they lifted any of the cylinders that the experiment was concerned with judging weights. They were told that they would be asked to lift the two standard weights and to "try to remember how they feel that you can make your judgments later of the other weights." These *Os* were required to perform the same orienting task that the incidental learners performed, *i.e.* placing the weights in the circles with the colored figures.

The scores of 13 *Os* were eliminated on the basis of their poor performance on the second series of judgments. If *O* failed to judge as many as five of the seven test-weights correctly in the second series, his judgments on the first series were not included in the calculations. Seven of the 13 *Os* who were eliminated were from the group of intentional learners and 6 were from the group of incidental learners. Among the intentional learners two had received one learning trial, two had received three, and three had received five learning trials; among the incidental learners, three had received one trial, two had received three, and one had received

five trials. The distribution of scores on the second series of judgments for the 120 acceptable *O*s was as follows: 61 of the *O*s, 31 intentional and 30 incidental learners, had all seven test-weights correct; 38 of the *O*s, 19 intentional and 19 incidental learners, had six of the weights correct; and 21 of the *O*s, 10 intentional and 11 incidental learners, had five of the test weights correct. Since approximately 90% of all the *O*s scored five or more correct on the second series and since the distributions for the incidental and intentional learners were very similar, *E* felt justified in his elimination of the 13 deviant *O*s.

Experiment II. After the first experiment was completed, Experiment II was conducted. The second experiment duplicated the first experiment with the following exceptions: *O* was explicitly informed that he might judge the seven test-weights to be almost equally distributed among the three categories, "lighter," "heavier," or "in between," or he might assign them to only two categories, or, perhaps, even to only one. Furthermore, *O* was cautioned not to try to guess where a weight belonged before lifting the weight.³

Results: Experiment I. The mean number of correct judgments for each of the six subgroups in the first experiment is shown in Table I. An analysis of variance

TABLE I
MEAN NUMBER OF CORRECT JUDGMENTS IN THE FIRST SERIES OF EACH EXPERIMENT

No. of learning trials	Experiment I		Experiment II	
	Intentional	Incidental	Intentional	Incidental
1	4.95	4.10	5.90	5.60
3	5.50	4.90	6.25	6.35
5	5.80	5.25	6.30	5.85

yielded an *F* significant at the 1% level for the difference between incidental and intentional learning and for the difference in the number of learning trials. The interaction was not statistically significant.

Two explanations of this unexpected superiority of intentional learning are available. First, although intertrial rehearsal was successfully eliminated by the use of our weight-lifting task, intratrial rehearsal probably was not. Many of the intentional learners, despite instructions to the contrary, were observed to be shaking or jiggling the standard weights as these weights were being held or placed in the circles. None of the incidental learners was ever observed to do this. The superiority of intentional learning, then, may be accounted for in terms of differential intratrial rehearsal.

The other explanation of the superior performance of the intentional learners is that they had greater confidence in their judgments and were more resistant to misleading instructions than the incidental learners. It will be recalled that all seven of the test-weights fell between the two standards. The instructions to the *O*s, however, undoubtedly suggested that one or more of the test-weights would be

³ The number of *O*s used was 127. Seven of the *O*s, four incidental and three intentional, were eliminated for failure to satisfy the criterion of five or more correct judgments on the second series.

heavier than the heavy standard and that one or more would be lighter than the light one. As a matter of fact, many of the *O*s who made errors in their judgments, intentional as well as incidental learners, commented, after the experiment had terminated, to the effect that they "really thought" that the weights all belonged "in between," but that they had not called them all that way because they thought some of the weights were supposed to be lighter or heavier. The superiority of intentional learning, therefore, may indicate, not that the intentional learners learned more or better than the incidental learners, but merely that they were more confident of what they had learned and consequently were influenced to a lesser degree by the misleading instructions than were the incidental learners.

Experiment II. To determine whether it was a lack of confidence and the consequent greater susceptibility to misleading instructions on the part of the incidental learners that resulted in their comparatively poor performance, Experiment II was performed. *E* believed that the modified instructions used in Experiment II, while not suggesting directly that all of the weights should be assigned to one category, nevertheless made it less likely that *O* would assume that all three categories had to be used. It was reasoned that, if the superiority of intentional learning in Experiment I was in no way a consequence of differential confidence and suggestibility, then the same results would be found in Experiment II as in Experiment I. If the intentional learners' superiority was due, at least in some degree, to this variable, then the differences between incidental and intentional learning in this second experiment, would be smaller.

The mean number of correct responses on the first series of judgments for each of the subgroups is shown in Table I. None of the *F*-values obtained in an analysis of variance was reliable. Not only were the incidental learning scores found to be as high as the intentional learning scores, but the scores after only one trial were not significantly lower than the scores after three or five trials. The incidental learners did perform as well as the intentional learners when their greater suggestibility was not used to mislead them.

Conclusions. Taking the results of both experiments into consideration, the following conclusions may be drawn: (1) under both incidental and intentional conditions of learning, *O*s can learn the weights of cylinders merely by lifting them; (2) the weight-lifting task, while it may successfully eliminate the opportunity for intertrial rehearsal, does not eliminate intratrial rehearsal by the intentional learners; (3) incidental learners are less confident that they remember what they have learned and are less resistant to misleading instructions than intentional learners; and (4) resistance to misleading instructions increases and, as a consequence, performance improves with increases in the number of learning trials.

THE SERIAL-POSITION CURVE AS A FUNCTION OF ORGANIZATION

By JULIUS WISHNER, THOMAS E. SHIPLEY, JR., and
MARVIN S. HURVICH, University of Pennsylvania

The shape of the serial-position curve can be affected by a variety of conditions, such as typographical emphases and patterning within the list.¹ One line of experimentation stems from Von Restorff's demonstration that isolated items are better recalled than massed items, a phenomenon interpreted in perceptual terms. The question has been raised, however, whether the finding could not be accounted for in terms of the different numbers of items of the massed and isolated types. Siegel demonstrated that the Restorff-effect holds for relatively isolated items when the absolute number of isolated and massed items is the same.² Saul and Osgood failed to confirm these findings in all respects,³ but they were confirmed by Postman and Phillips for intentional learning.⁴

Analysis of these experiments may offer a clue to some of the factors involved in the Restorff-effect. The fact that it is produced only with intentional learning, and not with incidental learning, would seem to indicate that it is not merely the objective organization of the list which produces the effect, but that S's set may also be an important variable. In other words, a relatively complex scheme for the organization of the materials is not necessarily transmitted to S automatically, i.e. by the mere fact that the presentation of the materials is based on some principle. It may well be that the failure of Saul and Osgood to produce the same results as Siegel, and Postman and Phillips, is a function of their instruction that a series of nonsense-syllables and numbers would be presented in "random order." Obviously, such information is misleading, and may well set S against any attempt to organize the material. (Siegel does not report giving such information to his Ss.) It is interesting to note in this connection Köhler's statement of the principle of memory as a function of organization: "intentional memorizing amounts to intentional organizing."⁵

In the present study, isolation is produced without variation in type of material, and S's set with respect to the organization of the material is facilitated by instruction. For example, an experimental group is shown a list of nonsense-syllables, the first half of which is in one color and typography, and the second half in

* Received for publication August 14, 1956. The authors are indebted to Professor F. W. Irwin for advice in the planning and execution of this study, and to Mr. William Rosenblith for assistance in collecting the data of Experiment III.

¹ J. A. McGeogh and A. L. Irion, *The Psychology of Human Learning*, 2nd ed., 1952, 115-134.

² P. S. Siegel, Structure effects within a memory series, *J. exp. Psychol.*, 33, 1943, 311-316.

³ E. V. Saul and C. E. Osgood, Perceptual organization of materials as a factor influencing ease of learning and degree of retention, *ibid.*, 40, 1950, 372-379.

⁴ Leo Postman and Laura W. Phillips, Studies in incidental learning: I. The effects of crowding and isolation, *ibid.*, 48, 1954, 48-56.

⁵ Saul and Osgood, *op. cit.*, 373.

⁶ Wolfgang Köhler, *Gestalt Psychology*, 1947, 262.

another color and typography. *S* is instructed that he is to learn two lists simultaneously, but the actual manner of the presentation of the material is the same as it would be for one list. Under such conditions, Köhler's theory would seem to require that serial-position curves of such groups would contrast with the usual curve for a control group in such a manner as to conform with the perceptual organization of the list and the explicit set of *S*. In brief, lists organized into two parts should produce the equivalent of two normally bowed curves.

Method: (a) Experiments I and II. To the control groups, a list of nonsense-syllables (33% association-value according to Glaze) was presented by means of a memory-drum at a 4-sec. rate of exposure. The syllables were so rotated that each *S* had a list beginning with a different syllable. The method of serial anticipation was used, and *Ss* were instructed accordingly. Learning was to a criterion of one perfect recitation. The same procedure was used for the experimental groups, except that the first seven syllables were typed in black caps and the second seven syllables were typed in red lower-case letters. *S* was informed that *E* was trying to discover how people learn two lists simultaneously and that the two lists were distinguished by differences in color and type. Experiment II was performed because it was thought that the irregularity in the curve for the control group in Experiment I might be related to the presence of a unique syllable (*LEX*) in the list. Another syllable was therefore substituted and the experiment replicated. The results in both experiments were alike in all essential respects and have therefore been combined. The 56 *Ss* were undergraduate women, 28 in the combined control group and 28 in the combined experimental group.

(b) Experiment III. A list of 18 nonsense-syllables (60% association-value according to Glaze), was presented at a 2-sec. rate of exposure to the criterion of one perfect recitation. The list was homogeneous for the control group of 18 undergraduate *Ss*, while for the experimental group (18 *Ss*) the first 6 syllables were typed in green caps, the second six in black lower-case letters, and the last 6 in red caps. All other conditions were like those of the previous experiments.

Results and discussion. The results are presented graphically in Fig. 1. It is clear, first of all, that the shapes of the serial-position curves are quite different for the control and experimental groups. The control groups showed singly bowed curves of the customary variety, while the experimental groups showed multiply bowed curves corresponding to the patterns of differentiation encountered. The second main finding is that learning is facilitated by differentiation. For the 14-syllable lists, the mean number of trials to the criterion was 18.7 for the experimental group and 22.8 for the control, while for the 18-syllable list, the corresponding values were 20.0 and 31.4. Both differences were significant at well beyond the 5% level.

In the third experiment, an analysis was made of intrusions. As might be expected, the great majority of intrusions were of the anticipatory type, *i.e.* syllables following the one to be anticipated were given. Some trend was noted, however, toward intrusions based on serial position as between the sub-lists in the experimental group. Four *Ss* gave the syllable in objective Position 5 in their anticipations of Syllables 17 a total of 11 times; *i.e.* these *Ss* seemed to be intruding the next

to the last syllable in what was for them the first 'list' into the position filled by the next to the last syllable in what was for them the third 'list.' This kind of error occurs a sufficient number of times to be quite noticeable in a chart of the intrusion-data, and would seem to indicate that Ss accepted the instruction to organize the list into three parts.

In the present series of experiments, both the objective organization of the list

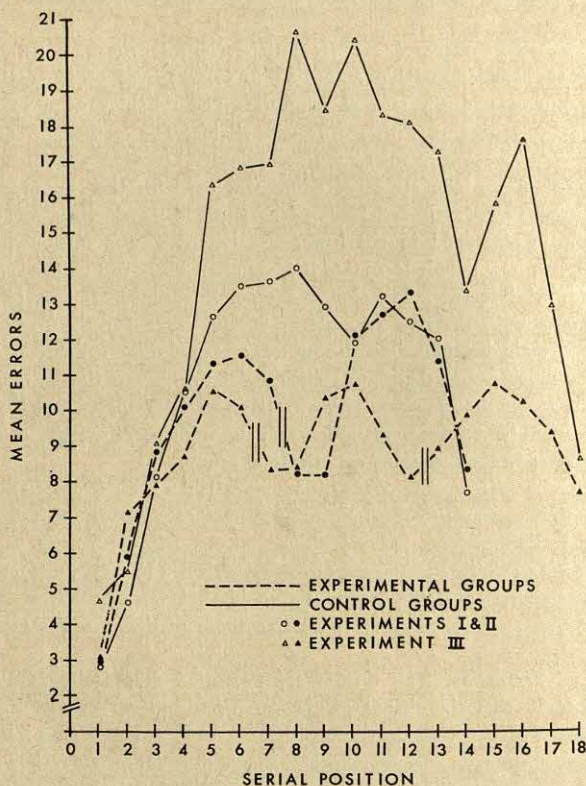


FIG. 1. SERIAL-POSITION CURVES

$N = 28$ in Experiments I and II combined and 18 in Experiment III. The paired vertical lines on the experimental curves indicate the points at which color and typography of the nonsense-syllables changed.

as well as S's set were manipulated simultaneously. It is evident that our instructions to S tended to facilitate the partition of the list by the experimental groups, although partition might have occurred on the basis of the objective differentiation alone. The experimental procedure chosen does not permit an analysis of the influence of each of these variables separately. This question is irrelevant, however, if one is interested in the effects of an explicit set to organize on the shape of the serial-position curve; for, as we have pointed out in our analysis of the Saul

and Osgood study, to leave this to the whim of *S* is to lose control over what may be an important variable involved in his performance.

The inference that the experimental *Ss* actually grouped the stimulus-material seems supported by the finding that they learned their lists in fewer trials than the control *Ss*. An analogy is suggested to the study by Mathews, who found that memory is facilitated by meaningful grouping of the stimulus-material.⁷

The degree to which intentional organization is a factor in serial learning is further indicated by the apparently systematic irregularities observed in the curves of the control groups. In Experiments I and II, these occur around the tenth position. In Experiment III, a significant dip occurs at Position 14. It is necessary to refer at this point to the absolute length of the list used in these studies, as compared to the length of the list used in most studies of the serial-position curve. Both Hovland and Hull have used lists longer than 12 syllables. In Hovland's study, with a 14-syllable list, the first descent of the curve is also at Position 10, although there is no subsequent rise, as in our data.⁸ In Hull's study, with a 15-syllable list, there is a descent at Position 9, and a subsequent rise.⁹ In our data, particularly those of Experiment III, there is a significant difference between the mean number of errors at Position 14, as compared with both Positions 13 and 15 for the control group, which might be attributed to some attempt by *S* to partition the list for easier learning when its length becomes too great to memorize as a whole. This sort of dip in the curve might not occur in a 12-syllable list, and might not be as apparent even in a 14- or 15-syllable list, but is boldly revealed by an 18-syllable list. It should be noted that similar rates of exposure were used in the three studies being compared. Our instructions may facilitate spontaneous organization if only because they do not specifically forbid it. On the other hand, Hull's typical instructions specifically request *S* to be passive.¹⁰ In the case of Experiments I and II, the longer exposure-time (4 sec.) may be an additional factor responsible for the dip in the curve for the control group, in that more time is available to *S* for organization.

These data suggest, therefore, that *S*'s approach to the problem of memorizing a list of nonsense syllables consists first of all in the identification of the terminal points of the list. These points serve as anchors for the memorization of the rest of the list. Particularly relevant here is a study by Workman, in which knowledge of where a list begins and ends was hidden from *S*.¹¹ Concealment was achieved by omitting some syllables from the list after the first presentation. The intent was to separate "the effect of primacy and ultimacy of the stimuli from the effect of perceptual primacy and ultimacy."¹² Workman's plot of the mean number of

⁷ Ravenna Mathews, Recall as a function of number of classificatory categories, *J. exp. Psychol.*, 47, 1954, 241-247.

⁸ C. I. Hovland, Experimental studies in rote-learning theory: VII. Distribution of practice with varying lengths of lists, *ibid.*, 27, 1940, 271-284.

⁹ C. L. Hull, The influence of caffeine and other factors on certain phenomena of rote learning, *J. gen. Psychol.*, 13, 1935, 249-274.

¹⁰ C. L. Hull, C. I. Hovland, R. T. Ross, M. Hall, O. T. Perkins, and F. B. Fitch, *Mathematico-Deductive Theory of Rote Learning*, 1940, 16.

¹¹ W. G. Workman, An experimental investigation of cognitive factors as contrasted with noncognitive factors in rote serial learning, Unpublished doctoral dissertation, University of Chicago, 1951.

¹² *Ibid.*, 9.

errors against the ordinal position of syllables in the list yielded a nondescript curve, somewhat resembling a double bow with a tail added at the end. On the other hand, a more or less classically bowed serial-position curve was obtained when the mean number of errors was plotted against the position of the items in the list as perceived by *S*. The latter curve was determined on the basis of *S*'s report, and again by *E*'s judgments of *S*'s behavior, which correlated highly with *S*'s report. Thus, when the identification of the objective terminal-points is impossible, as in Workman's procedure, or when the length of the list is greater than the convenient memory-span of *S*, as is apparently the case in Experiment III, arbitrary cut-off or anchoring points are constructed by the *S*.

Orbist's experiment provides some hints as to *S*'s procedure for the learning of the whole list.¹³ He studied the *GSRs* of *Ss* as they were learning a list of nonsense-syllables. His data show that the *GSRs* follow the bowed serial-position curve: *i.e.* at the beginning of a series of presentations, *GSRs* are greatest at the ends of the list; as these become capable of being anticipated correctly, the *GSRs* become greatest for the other syllables, beginning at the beginning and proceeding towards the end as each syllable is anticipated correctly. If one interprets the *GSR* as an indicator of attention and anticipation, as does Darrow,¹⁴ it would appear that *S* first attends to the ends of the list and then proceeds from the beginning toward the end, as would be expected from Köhler's principle.

Summary. The purpose of the present series of studies was to throw light on some of the conditions underlying the bowed serial-position curve. In Experiments I and II, the control groups were presented with a list of 14 nonsense-syllables; the experimental groups were presented with the same syllables, except that the first seven were typed in black caps and the second seven in red lower-case letters, and they were instructed that they were to learn two lists simultaneously. In Experiment III, a list of 18 nonsense-syllables was learned; it was subdivided perceptually into three parts for the experimental group, but not for the control. Under all conditions, differentiation of the lists facilitated learning. The serial-position curves were doubly or triply bowed for the experimental groups, but systematically bowed for the controls. Together with data adduced from other studies, the present findings were utilized in an account of serial learning which supposes that *S* attempts to find some anchoring points in a list. The findings constitute an extension of the Restorff-effect, with amount of 'isolated' material equated with the 'massed,' but with interfering sets minimized.

¹³ W. D. Obrist, Skin resistance and electroencephalographic changes associated with learning, *Summaries of Doctoral Dissertations*, Northwestern University, 18, 1950, 607-610.

¹⁴ C. W. Darrow, Neural mechanisms controlling the palmar galvanic skin reflex and palmar sweating, *Arch. Neurol. Psychiat.*, 37, 1937, 641-663.

CRITERIA IN CONCEPTUAL TRANSPOSITION

By GEORGE A. TALLAND, Massachusetts General Hospital

Intelligent adults can categorize in terms of absolute or relational criteria, and can shift from one to the other as the situation requires. It would appear that they are capable of perceiving and of discriminating in both absolute and relative properties. Yet a recent experiment by Buss on conceptual transposition has provided results in which its author finds support for "a reinforcement-inhibition view rather than a relational view" of discriminative learning in college students.¹ The purpose of the present study was (a) to repeat Buss's work under more stringent control, and (b) to analyze further the conditions making for the selection of absolute or relational criteria. The theoretical position adopted can be stated in terms of the following hypotheses: (1) Given a universe divided into two distinct classes, *i.e.* two inclusive categories, a relational criterion of discrimination will be adopted in preference to an absolute one. The differential availability of a relational criterion can be further increased by embedding appropriate linguistic cues in the definition of the categories. (2) When a concept is formed within an exclusive category, a relational criterion will not be more readily available than an absolute one. (3) Relational inclusive categories are established more easily and absolute inclusive categories less easily than exclusive categories of either type.

Buss's experimental material and procedure were applied under conditions devised for the present experiment as well as those originally designed by him. To provide for more rigorous criteria of learning and to forestall chance-achievement, his procedure was modified in certain details. The measure of performance used was trials to criterion, since the aim was to specify availability or ease, the constructs involved in the three hypotheses to be tested.

Subjects. Seventy-two university students served as Ss. They were randomly assigned to five experimental groups. Twelve Ss were substitutes for others who failed to meet the standard of performance set for qualification.

Materials. The materials consisted of 120 blocks, of which 60 were 1 in. high, 30 were 0.5 in. high, and 30 were 2 in. high. They corresponded to Buss's specifications: black, blue, gray, green, red, and white, four round with surface-areas of 1.49, 0.68, 0.31 and 0.11 sq. in., three triangular and three square with surfaces of 1.56, 0.56 and 0.25 sq. in. Colors, forms, and sizes were so distributed within each class of height that no two blocks were identical in all four attributes.

Procedure. S faced E across a table on which a screen was placed to conceal the box containing the blocks. Instructions were standardized, but differed in one detail from one experimental group to another.

* Received for publication August 18, 1956. This experiment was conducted in the laboratories of the Department of Social Relations of Harvard University. The author wishes to thank Dr. J. S. Bruner for his assistance.

¹ A. H. Buss, Rigidity as a function of absolute and relational shifts in the learning of successive discrimination, *J. exp. Psychol.*, 45, 1953, 153-156.

For one they read: "You will be shown a series of colored blocks, one at a time. Some of these blocks are *VEC*, others are *BIX*." [Both nonsense words were then spelled.] "You do not know what *VEC* and *BIX* mean, and your task is precisely this: to discover what a *VEC* is and what *BIX* is. You will do this by calling out '*VEC*' or '*BIX*,' and only one or the other of these words, each time I show you a block. If you are right, I shall say 'right,' and if you are wrong I shall say nothing. Silence therefore means that your guess was wrong. All blocks you will see are either *VEC* or *BIX*. In this way you will be able to discover what is a *VEC* and what is a *BIX*." [For another group the two nonsense-words were *BIX* and *BIXER*.] For two further groups the instructions were: "You will be shown a series of colored blocks, one at a time. Some of the blocks are *VEC*, others are not *VEC*." [*VEC* was spelled.] "You do not know what *VEC* means. . . . Each time you are shown a block you either call '*VEC*' or '*not VEC*.'"

When *S* had clearly grasped his instructions, the blocks were shown singly in front of the screen, always with the criterial attribute (height) occupying the vertical dimension. Exposure-time was not standardized, but never exceeded 15 sec.

TABLE I

CATEGORIES OF DISCRIMINATION AND MEAN TRIALS TO CRITERION IN EACH SERIES FOR THE FIVE GROUPS

Group	Series 1		Series 2		Trials	
	2-in. blk.	1-in. blk.	1-in. blk.	0.5-in. blk.	Series 1	Series 2
A	VEC	BIX	VEC	BIX	6.25	8.00
B	VEC	BIX	BIX	VEC	6.42	30.50
C	VEC	not VEC	VEC	not VEC	8.75	17.42
D	VEC	not VEC	not VEC	VEC	6.42	18.25
E	BIXER	BIX	BIXER	BIX	7.00	9.08

After display, each piece was placed behind the screen; *S* never saw more than one block at a time. The first series consisted of 2-in. and 1-in. blocks for all groups, and terminated with six consecutive correct responses or with the sixtieth exposure. It was followed without break or warning by the first 0.5-in. block of the second series. The second series consisted only of 0.5-in. and 1-in. blocks, and extended at least to 32 presentations, or as many more as were needed to complete 6 consecutive correct responses, but not more than a total of 60. Order of presentation was so standardized that no consecutive 5 blocks could be correctly identified otherwise than by the appropriate principle of categorization, or by some unpredictable combination of complex criteria and chance. Sequences within each category, *i.e.* blocks of the same height, were limited to 3 in a run. *E* followed his scoring sheet while presenting the blocks. If *S* noticed the transposition and asked whether the same principle continued to hold, *E* answered affirmatively; any other remark or question was left unanswered. At the end of the experiment *S* was asked to explain what a *VEC* was or what distinguished a *BIX* from a *VEC* (or from a *BIXER*).

Design. There were five experimental groups (Table I). Groups A and B were instructed to discover how to discriminate *VEC* from *BIX*. For both, *VEC* was the correct response to 2-in. blocks and *BIX* to 1-in. blocks in Series 1. For Group A, the relation was transposed in Series 2, while for Group B, *BIX* remained the correct response to 1-in. blocks throughout and (no third class being available) *VEC* applied both to the tallest (Series 1) and shortest (Series 2) pieces. Group C and D

were like Groups A and B except that for them only one category could be defined. Group E received instructions which were intended to set the Ss for a relational criterion, the two category-labels being distinguished by the comparative adjectival suffix *-er*, and for this group a relational criterion was the correct one.

The score used in statistical analysis was number of blocks displayed prior to the six consecutive correct responses which constituted the criterion in each series. Hypotheses were tested by determining the significance of differences between group means. Since none of the arrays was normally distributed, Wilcoxon's non-parametric test for unpaired replicates was used.² If the hypothesis implied confirmation of the null hypothesis, a *t*-test or an analysis of variance also was made in order to reduce the probability of a Type II error. Differences among Groups A, B, C, and D (one vs. two categories, relational vs. absolute transposition, and interaction) were studied in an analysis of covariance relating performance after transposition to performance in Series 1. For these tests requiring normal distribution, scores included the six correct responses and were transformed by the square-root method.

Results. All Ss in Group E solved the discriminative problems of Series 1 in 30 trials and successfully transposed the principle to Series 2. One S instructed under Conditions A-B and two under Conditions C-D in Series 1 failed to reach the criterion in 60 responses, and a further S in each of Groups B, C, and D, failed in 30 responses. Since the scores of these last three exceeded their sample means by two SD, they were replaced by other Ss, as were also two Ss, each in Groups C and D, and one each in Groups A and B, who failed to reach criterion in Series 2. Mean number of responses preceding 60 correct calls in Series 1 and 2 are shown in Table I.

Differences among scores of Series 1 are not significant at the 5% level of confidence, whether tested by analysis of variance for ranked data,³ or for normally distributed data. In Series 2, the Wilcoxon test shows that the mean rank of Group B differs significantly from those of Groups A and E at the 1% level. Differences between Groups A and E, and between C and D, are not significant. Analysis of covariance points to a significant interaction effect between inclusive and relational categorization, as well as to the more ready transposition of the relational criterion as compared with the absolute, but shows no significant difference due to inclusive vs. exclusive categorization.

Discussion. The experiment confirms the prediction that under conditions of inclusive categorization the choice of an absolute or relational criterion is not a matter of chance, but that the latter is more likely to be adopted. It also supports the reasoning that, since it is easier to distinguish one class of objects from another than from *any* other class, and since a relational criterion is the most economical in a dichotomous universe, a relational inclusive category would be the most easily available—and, as a corollary, that an absolute inclusive category, being the least economical, would be the least likely to be chosen. The hypothesis that the

² Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures*, 1949, 4-5.

³ W. H. Kruskal and A. W. Wallis, Use of ranks in one-criterion variance analysis, *J. Amer. statist. Assoc.*, 47, 1952, 583-621.

selection of a relational criterion for inclusive categorization could be induced with greater probability by means of appropriate linguistic cues is not confirmed.

The task employed allowed for the adoption of a relational or an absolute criterion, whether categorization proceeded by establishing two distinct classes, or by establishing one class and excluding from it all objects which did not possess its defining attribute. Having completed this task of categorization, *S* had to solve a second problem, in which transfer of the first criterion could either help or hinder, according to the experimental design. Immediate success after transposition indicated that *S* had selected the criterion of classification common to both series of presentation, and the longer solution of the problem in Series 2 was delayed, the less likely it appeared that *S* had chosen or even considered this criterion. Inferring to cognitive strategy from the number of trials needed to achieve success, the results of the experiment suggest that absolute and relational criteria both are available for discrimination and categorization, though not equally so in any situation. Given the conditions of exclusive categorization in this experiment, the probabilities of adopting one or the other criterion are not significantly different; given the conditions of inclusive categorization, the relational criterion is distinctly more readily available, so much so indeed that an attempt to induce it by linguistic cues has no additional effect. Taking all conditions of discrimination and categorization sampled by this experiment, the probability of choosing a relational criterion is higher than that of an absolute one. This finding is attributed to the greater ease with which inclusive categories are established, and to the higher probability of formulating them in relational terms.

All these findings fail to confirm Buss's argument for primary stimulus-generalization based on S-R principles and agree with cognitive theory, which accounts for the adoption of an absolute criterion for concept-formation in terms of more complex processes. That the task was one of concept-formation rather than simple discrimination followed from the experimental design which varied irrelevant as well as criterial attributes of the objects presented from trial to trial. In all probability *Ss* considered both absolute and relational criteria; under certain conditions one type was more readily available and subsequently became confirmed or reinforced. Experiments with human *Ss* have one unquestionable advantage over animal experiments: *E* can ask the *Ss* what they thought they were doing. In this experiment, a large proportion of *Ss*, who might have appeared to adopt and transfer an absolute response, said, in fact, that *they had first learned a relational discrimination and later reversed it*. Some others who appeared to have chosen the correct relational criterion believed they had followed some such more complex relational principle as "blocks that are taller than wide." In fact, strict adherence to this type of criterion would not have enabled them to solve the problem in either series. These reports are, nevertheless, interesting for two reasons: (a) they point to a propensity to discriminate and categorize not simply in relational terms, but to formulate more difficult relational hypotheses than necessary, and (b) they suggest the operation of a perceptual or cognitive mechanism which aims at unity where there is discontinuity; rather than transfer a relationship from one situation to another, an attempt is made to subsume both under a relational principle which need not be shifted.

Summary. Five groups of Ss, each working under different conditions, were set the task of learning a discrimination which entailed concept-formation, and to transfer the concept to another discriminative problem. Three groups could achieve this task by transposing a relational criterion, two groups by transferring the concept of an absolute stimulus. There was no break or other indication of some change between the two situations. Concept-formation proceeded for two groups by defining the criterial attribute of one category and excluding from it all objects not thus defined. Two groups defined both categories by their criterial attribute, and a fifth group did the same, but was also given a verbal cue intended to suggest a relational hypothesis.

The Ss who had to categorize inclusively tended to do so relationally, and to transpose the relational solution. This tendency was not increased by the additional linguistic cue—presumably because it was already sufficiently strong. With exclusive categorization, relational criteria did not seem to be more readily available than absolute ones.

The experiment supports the argument that the construct of primary stimulus-generalization is inadequate for analyzing conceptual processes. When solving problems of categorization, intelligent adults are able to formulate hypotheses both in relational and in absolute terms, and will test either. In certain instances of inclusive categorization, the relational criterion appears to be the more readily available, and these are the conditions under which correct concepts are formed most readily.

ON THE VALIDITY OF THE POINT-ASSIGNMENT PROCEDURE IN THE CONSTANT-SUM METHOD

By FRANK J. DUDEK and KATHERINE E. BAKER, University of Nebraska

Scales for psychological variables have been found to vary with a number of experimental conditions or specific features of scaling method.¹ Another important question for investigation concerns the extent to which particular procedures used for reporting judgments may themselves affect the scales.² The constant-sum method has many advantages and provides a sound rational approach to scaling, but previous studies suggest that the manner of indicating perceived ratios may introduce an artifact into resultant scales.³ With this method of reporting perceived ratios, different biases appear to be operating when the stimuli in the pair are similar to each other than when they are quite dissimilar.⁴ While such biases may be valid reflections of perceptual effects of similarity upon perceived ratios, certain characteristics and restrictions of the method might provide an alternate explanation. It seems possible that *O* may report point-assignments without adequate recognition or appreciation of the lack of linearity between change in point-assignment and change in implied ratios. Even with direct reporting of ratios, one may still ask whether *O* really appreciates what a judgment of 10-1 means in terms of the perceptual experiences involved.⁵ What is needed to answer the question is a comparison of results from a method utilizing point-assignments and from a method which allows expression of perceived ratios without use of numerical values. In the study reported here, *O* expressed perceived ratios by setting a pointer to divide a line into two parts bearing the same relation to each other as the members of the pair of stimuli. Each *O* also made settings of the pointer to represent certain numerically stated ratios, thus providing a check on the ratio which he wished to imply by his line division. Since it was possible to employ *O*s who had judged the same stimuli in terms of point divisions about three years previously, information was obtained on the consistency of judgment over a relatively long period of time and with two quite different ways of reporting.

* Received for publication May 4, 1956.

¹ See, for example, Harry Helson, Adaptation-level as a frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29; W. R. Garner, Context effects and the validity of loudness scales, *J. exp. Psychol.*, 48, 1954, 218-224; K. E. Baker and F. J. Dudek, Weight scales from ratio judgments and comparisons of existent weight scales, *J. exp. Psychol.*, 50, 1955, 293-308; Trygg Engen, An evaluation of a method for developing ratio scales, this JOURNAL, 69, 1956, 92-95.

² Since the completion of this study, two articles concerned with some of these issues have been published. For relevant data on scaling loudness, see S. S. Stevens, The direct estimation of sensory magnitudes—loudness, this JOURNAL, 69, 1956, 18-19, 24-25. See also W. R. Garner, H. W. Hake, and C. W. Eriksen, Operationism and the concept of perception, *Psychol. Rev.*, 63, 1956, 155-157, for an independent parallel analysis of the scaling-problems to which the results reported here are directly relevant.

³ Baker and Dudek, *op. cit.*, 304-306.

⁴ *Ibid.*, 300-304.

⁵ *Ibid.*, 300-304.

Procedure. The apparatus was essentially the same as previously described, and the same nine weight-stimuli were employed.⁶ A device was added to the apparatus with which *O*, by turning a crank that moved a pointer along a line, could divide a 25-in. line into two segments which would indicate the relative magnitudes of the paired weights. At the back of the apparatus, not visible to *O*, was a scale dividing the line into 100 units by means of which *E* could translate the line-division directly into equivalent point-estimates.

Four graduates students, who were members of a group that had served in a scaling experiment three years previously, served as *Os*. Four experimental sessions

TABLE I

SCALE-VALUES FOR WEIGHTS AND DIVISIONS OF A LINE INTO DESIGNATED RATIOS

Scale-values for weight-stimuli from physical measurement and from group and individual judgments with relative magnitude indicated by line-division (*LD*) and by point-assignment (*PA*). *A* = 1.00 is the unit of each scale.

Stimulus with physical measurement								
<i>Os</i>		<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>I</i>
		1.66	2.30	3.62	3.93	5.24	6.36	8.48
<i>SA</i>	<i>LD</i>	1.37	1.64	2.35	2.55	3.46	3.94	5.19
	<i>PA</i>	1.25	1.79	2.21	2.51	3.07	3.91	5.50
<i>BS</i>	<i>LD</i>	1.36	1.80	2.47	2.69	3.64	4.39	4.01
	<i>PA</i>	1.32	1.64	2.41	2.60	3.28	3.87	5.39
<i>CK</i>	<i>LD</i>	1.98	2.42	4.18	4.59	6.55	8.06	12.54
	<i>PA</i>	1.49	2.36	4.10	4.55	6.27	8.31	12.12
<i>WW</i>	<i>LD</i>	1.32	1.62	2.06	2.15	2.55	3.02	3.73
	<i>PA</i>	1.34	1.79	2.32	2.64	3.27	4.18	5.36
Group	<i>LD</i>	1.44	1.80	2.56	2.76	3.65	4.34	5.77
	<i>PA</i>	1.34	1.81	2.60	2.90	3.70	4.65	6.37

Ratios between smaller and larger segments of a line when *O* was asked to divide it in accordance with designated ratios.

Designated ratio							
<i>Os</i>	1:1	1:2	1:3	1:4	1:6	1:8	1:10
<i>SA</i>	1.06	1.99	2.88	3.76	5.56	8.64	12.56
<i>BS</i>	1.03	1.91	2.92	3.82	5.06	7.25	10.43
<i>CK</i>	1.04	2.14	3.14	4.30	5.90	9.53	14.69
<i>WW</i>	1.01	2.23	3.19	4.20	5.90	8.64	10.43
Group	1.03	2.06	3.03	4.01	5.58	8.44	11.80

provided four judgments of each pair of stimuli (two judgments with the heavier weight lifted first and two judgments with the lighter weight lifted first). The order of presentation of paired stimuli was identical to that used in the earlier experiment. (At the end of each session of weight-judgments, *O* was asked so to divide the line that the two parts appeared to represent the relative magnitudes of 1:1, 2:1, 3:1, 4:1, 6:1, 8:1, and 10:1. Two settings for each ratio were made in counterbalanced order, with the larger line segment once to *O*'s left and once to his right.) Scale-values for the weight-stimuli were computed from the line-division data for each *O* individually and for the pooled group data. Comparable

⁶ *Ibid.*, 294.

scale-values were computed for these *O*s from their point-assignment judgments in the earlier experiment.

Results. Scale-values for the two procedures of indicating judgments are presented in Table I, which includes also the results of the line-division settings for the

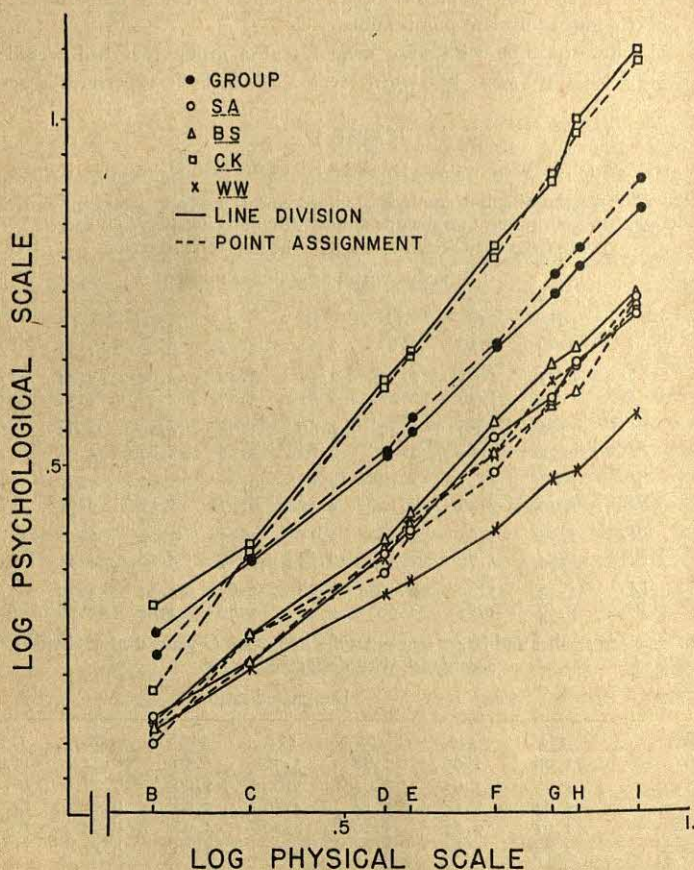


FIG. 1. RELATION BETWEEN PSYCHOLOGICAL AND PHYSICAL SCALE-VALUES FOR INDIVIDUAL *O*s AND FOR THE GROUP WITH MAGNITUDES REPORTED IN TERMS OF LINE-DIVISION AND IN TERMS OF POINT-ASSIGNMENTS

The group curves have been displaced upward one-tenth of a log-unit to avoid overlap with the individual curves. $A = 1.00$ is the unit of each scale.

indicated ratios. The group-scales under the two conditions manifest a high degree of correspondence, which is especially striking since they represent judgments made three years apart. It is clear that there are consistent individual differences. Some *O*s (like *CK*, for example) expand the scale, while others contract it with reference to the physical scale. At the same time, it is important to note that relations among stimuli remain quite stable from scale to scale. Thus, all scales indicate that the

ratios E/D and H/G are about the same, and that these are different, for example, from the ratios G/F and I/H which are also similar.

That the results are apparently not attributable to lack of appreciation of what 80-20 means is indicated by reference to the data for line-division settings of the specified ratios. For the group, the division of the line into perceived ratios closely approximated physical ratios up to 4:1. For greater ratios (8:1 and 10:1), the tendency was to make the difference in line-segments greater than physical measurement would indicate. This tendency was fairly uniform for all O s represented.

The consistencies of the scales are clearly revealed when the psychological values are related to the physical scale-values in a log-log plot, as shown in Fig. 1. A log-log plot makes proportional differences at all scale-values appear equal. The scales are very much alike whether relative magnitude was indicated by dividing 100 points between stimuli or by line-division. Here the expansion of the scale by CK , a consistent result for both methods of indicating judgments, is clearly seen. SA and BS also demonstrated striking consistency, while WW was somewhat less consistent.

Another way to consider O and scale-consistency is in terms of correlations between different trials and conditions. For the data involving line-division, the test-retest coefficients for the same pairs of weights judged in two different sessions are 0.885, 0.876, 0.868, and 0.718 for SA , BS , CK , and WW , respectively. When the averages of two trials involving the same pair are correlated, the reliabilities are 0.95, 0.94, 0.92, and 0.82, respectively. This increase in reliability corresponds very closely to that predicted from the Spearman-Brown prophecy formula. When the average of two trials for judgments in terms of point-assignment is correlated with the average for two trials (three years later) for judgments by line-division, the reliabilities are 0.91, 0.91, 0.92, and 0.86, respectively. The latter might be described as alternate-form reliabilities with a test-retest interval of three years. For the line-division data, where the average for two O s (two trials each) is correlated with averages from two other O s (two trials each), the correlation is 0.95; and when the averages of single judgments by each of four O s is correlated with another judgment by the same four O s, the coefficient is 0.974. Correlating point-assignments from all four O s (two estimates for each pair) with line-division estimates from all four O s, the result is 0.96.

Discussion. These data require little comment. At least for sophisticated judges, scale-values obtained with the constant-sum method provide a basis for quite accurate prediction of scale-values derivable from a different method of expressing perceived ratios; namely, line-division. The assumption that point-assignments are valid reflections of the perceptions involved thus appears reasonable. The rather considerable individual differences in scales have important implications since it seems reasonable to expect that similar differences will be found for many dimensions. The usefulness of a scale derived for a group will be limited to the extent that the sample for which the scale is obtained represents the population for which the scale will be used.

Summary. An experiment was conducted to examine the possibility that reporting judgments in numerical terms might influence the resulting scales. Almost identical scales were obtained for lifted weights when perceptions were reported in numerical or non-numerical terms. For the O s represented, there was no evidence that the use of point-assignments introduced any distorting effects.

RESTRICTION OF RANGE IN THE JUDGMENT OF SINGLE STIMULI

By ALLEN PARDUCCI, Swarthmore College, and
RAYMOND HOHLE, University of Oregon

The change in absolute judgment which follows an extension of the range of stimuli judged tends to be much greater than that which follows restriction of the range.¹ Extension of the range may indicate to *O* that the stimuli belong to a new and wider distribution for which a modified scale of judgment is appropriate; restriction of the range may not indicate the need for a modified scale, since the missing values are expected to recur upon subsequent trials. In an experiment designed to evaluate this interpretation, *Os* were provided with different kinds of information about the permanence of the restriction of range.

Method: Observers. Students of introductory psychology were randomly assigned to 8 groups of 11 *Os* each.

Materials. The stimuli were circles evenly spaced on narrow pages (2 × 11 in.) which were stapled together to form booklets. The diameters of the circles varied from 0.25-1.25 in. in eight 0.125-in. steps (giving nine different sizes). Below each circle was a 0.5-in. horizontal line on which the judgment was to be recorded.

Procedure. Each *O* was given a booklet and a page of instructions which read as follows:

Your task in this experiment is to judge the sizes of the circles appearing in the response booklet. As you come to each circle, make a judgment of how large or small it appears to you. Record your judgment in terms of one of the following size categories: Very Large—*VL*; Large—*L*; Medium—*M*; Small—*S*; and Very Small—*VS*. Write the appropriate letter-abbreviation on the horizontal line under the circle being judged. Work through the booklet from beginning to end, judging each circle in the order of its appearance. Do not turn back to change any of your previous judgments. This experiment is concerned with your subjective impressions of each of the circles when you first see them; and since you are to judge how large or small they look to *you*, there can of course be no right or wrong answers. After you have read these instructions carefully, open your booklet to page one, and begin working.

The first 15 pages of each booklet contained the pre-shift circles. For all *Os*, there were 88 of these circles, 6 on pages 1-14, and 4 on page 15. For 4 of the groups (the *N*-groups) the circles of different sizes were distributed as though drawn from a negatively skewed population, the 5 largest circles appearing 16 times each and the 4 smallest 2 times each. The remaining 4 groups (the *P*-groups) judged a pre-shift set of circles whose sizes were distributed as though drawn from a positively skewed population, each of the 5 smallest circles appearing 16 times and each of the 4 largest 2 times. The order in which the different sizes occurred was quasi-random and identical to that for the corresponding groups of a previous

* Received for publication April 16, 1956.

¹ Allen Parducci, Direction of shift in the judgment of single stimuli, *J. exp. Psychol.*, 51, 1956, 169-178.

experiment.² For all groups, the 100 post-shift circles included only the 5 smallest sizes, each of which appeared 20 times in quasi-random order.

Two control groups (one *N*- and one *P*-group) proceeded directly to the post-shift circles with no break between the two series. The other six groups were provided with information concerning the possible future recurrence of circles to be omitted during the post-shift series. This information appeared on a separate page inserted in the booklets immediately following the pre-shift series. There were three alternative sets of information ('Yes,' 'Maybe,' and 'No') calculated to create three different kinds of expectancy. Each set appeared in the booklets presented to one of the *N*- and one of the *P*-groups. The information read as follows:

['Yes'] On the immediately following pages, some of the circles which you have been judging may appear to have been omitted. These circles will appear again on later pages.

['Maybe'] On the immediately following pages, several of the circles which you have been judging will be omitted. In some booklets, these circles will not appear again, in others, they will appear on later pages.

['No'] For the remainder of the booklet, several of the circles which you have been judging have been omitted. These omitted circles will not appear again in this experiment.

The hypothesis tested was that adjustment to the shift would be a function of the permanence of change in the stimulus-series suggested by the interpolated information.

Results. The judgments were converted to numerals: *VL*, 5; *L*, 4; *M*, 3; *S*, 2; *VS*, 1. To simplify tabulation, only the judgments of the middle stimulus, the largest circle of the restricted post-shift series, were used. Two measures of each *O*'s change in judgment were computed: (a) his initial shift, the mean of his first 3 post-shift judgments of the middle stimulus minus the mean of his last 3 pre-shift judgments of that stimulus; and (b) his total shift, the mean of his last 10 post-shift judgments of the middle stimulus minus the mean of his last 3 pre-shift judgments of that stimulus. Logarithmic transformations were made of all scores (after 2.0 had been added to each to eliminate negative values) in order to reduce heterogeneity of variance.

The group-means are presented in Table I. Information to the effect that the missing stimuli would not or might not recur increased the shift in judgment relative to the control in the *N*-groups. The shift was greater in *N*-*No*' than in *N*-*Maybe*, but there was no difference between *N*-*Yes*' and *N*-*Control*, an outcome which supports the interpretation that, in the absence of special information, *O* assumes that the missing stimuli will eventually recur.

The neat ordering of the *N*-groups was not, however, found in the *P*-groups. There was little difference among the informed *P*-groups, although all shifted more than the control. This difference in the importance of the information for the *N*- and *P*-groups is reflected in the significant ($p < 5\%$) interactions obtained in the analyses of variance. The analyses of the scores for both initial and total shift indicated that while the effects of the different sets of interpolated information were different from each other ($p < 1\%$), this difference depended upon the direction of

² *Ibid.*, 170.

skewness of the pre-shift series. In all cases, the shifts were much smaller for the *P*- than for the *N*-groups. (Variance due to skewness was significant beyond the 1% level.)

Discussion. The results for the groups shifted from the negatively-skewed series are consistent with the hypothesis that adjustment to the post-shift series would depend upon interpolated information regarding the permanence of change in the stimulus-series. They suggest that the direction-of-shift effect found in an earlier experiment depends upon *O*'s interpretation of the restriction as a temporary affair.³ The relatively large shift in judgment by the *P*-*'Yes'* group (informed that the missing stimuli would recur) is not, however, consistent with the hypothesis. It may be that for all informed *P*-groups the effect of the information was merely

TABLE I
MEAN SHIFT IN JUDGMENT
(All scores logarithmically transformed.)

Group	Initial		Total	
	Mean	SD	Mean	SD
<i>N</i> - <i>'No'</i>	1.380	0.250	1.456	0.133
<i>N</i> - <i>'Maybe'</i>	1.139	0.286	1.256	0.256
<i>N</i> - <i>'Yes'</i>	0.912	0.151	1.058	0.133
<i>N</i> -Control	0.914	0.217	1.031	0.197
<i>P</i> - <i>'No'</i>	0.823	0.159	0.861	0.215
<i>P</i> - <i>'Maybe'</i>	0.763	0.200	0.927	0.085
<i>P</i> - <i>'Yes'</i>	0.877	0.141	0.869	0.207
<i>P</i> -Control	0.591	0.240	0.664	0.240

to call attention to a change in the stimulus-series which would not otherwise have been noticed. Since the stimulus-shift involved the omission of stimuli which had occurred only infrequently during the pre-shift series, the meaning of the interpolated information may have been less clear for *P*-groups. By the time they realized that the largest circles were being omitted, they might well have forgotten the significant details of the interpolated information.

The smaller shift by the *P*-Control than by the *N*-Control corroborates previous findings with a different set of stimuli and a different method of presentation.⁴ It indicates that judgmental data cannot be completely described as a function of the stimulus-range, but that some account must be taken also of the distribution of stimulus-values within the range. The results of this experiment are consistent with the more general assumption that judgments are made as though the judged stimulus were being compared with the whole series of stimuli with which *O* expects to be presented. Events which determine what *O* expects will determine his judgments. Thus, a prior verbal description of the stimulus-series tends to minimize subsequent change in judgment,⁵ and information calculated to affect *O*'s interpretation of a stimulus-shift also will affect the magnitude of change.

³ *Ibid.*, 169-178.

⁴ *Ibid.*, 169-173.

⁵ Parducci, Learning variables in the judgment of single stimuli. *J. exp. Psychol.*, 48, 1954, 24-30.

Summary. This experiment explored the possibility that absence of judgmental shift following restriction of stimulus-range reflects *O*'s belief that the missing values will recur. Judgments of absolute size were made of a series of circles varying in diameter. The series began with a wide range of values and then shifted to a narrower range. Different groups of *O*s were provided with different sets of interpolated information concerning the permanence of the shift. This information had significant effects upon subsequent judgments. Shifts in judgment were greater when *O*s were informed that the stimulus-shift would be permanent. The role of information was, however, found to depend upon the nature of the stimulus-shift, the specific information being effective only when the shift was easily discriminated.

THE EFFECT OF TIME ON SIZE-CONSTANCY

By W. L. GULICK and R. E. STAKE, Princeton University

In the experimental study of the relation between size and distance, little consideration has been given to exposure-time as a determinant of size-estimates. In a recent study, however, Leibowitz, Chinetti, and Sidowski did include exposure-time as a variable in the investigation of size-constancy. They illuminated a hallway for intervals of 0.0005 sec. and found that estimates made under this condition were "essentially identical with those obtained under continuous illumination, demonstrating that reduced exposure-time has little effect on the perception of size."¹ In that experiment a fixation-point of light was presented at the standard before the standard was illuminated, and such other cues to distance as perspective and interposition were not eliminated during its illumination.

The present experiment was undertaken to determine what effect limited exposure-time might have on size-constancy under conditions similar to those reported by Chalmers² and by Holway and Boring.³ Cues to distance in the present investigation were limited to those of accommodation, convergence, and retinal disparity.

Observers. Each of 12 Os, aged 20-34 yr., met a visual acuity criterion of 1.0 (corrected vision) on a minimum-separability test. Each was naïve concerning the purpose of the experiment and the theoretical aspects of constancy phenomena.

Apparatus. The experiment was conducted in a darkened tunnel 130 ft. long. Five cut-out triangles were used as standards, one being placed at 20, 30, 40, 50, and 80 ft. from *O*. The altitude of each standard subtended a visual angle of 1°. By remote control, *E* could move any standard in and out of *O*'s field of vision. Concealed lighting on a white panel at the end of the tunnel produced a homogeneous reflectance of 3.0 ft.-L and provided illumination for the cut-out triangles. A comparison-triangle, geometrically similar to the standards, was set 20 ft. from *O* and to the left of the standards (3° separation, base mid-point to base mid-point). Illuminated to match the brightness of the white panel, the comparison was variable in altitude from 0° 12' to 4° 48'. The walls of the tunnel were so curtained that the triangles appeared in a totally darkened field.

An episcotister permitted exposures of 0.1 and 0.8 sec. For trials on which either of these two brief exposures was used, a manually operated shutter enabled *E* to limit *O* to single, rather than repeated, exposures. The shutter was used also to present exposures of 4.0 sec.

* Received for publication August 2, 1956. The authors are indebted to L. J. Baranski, D. P. Thomas, and W. M. Smith for their assistance in the planning and execution of this experiment; and to M. B. Wilk for his opinions regarding the analysis of the data.

¹H. Leibowitz, P. Chinetti, and J. Sidowski, Exposure duration as a variable in perceptual constancy, *Science*, 123, 1956, 668-669.

²E. L. Chalmers, Jr., Monocular and binocular cues in the perception of size and distance, this JOURNAL, 65, 1952, 415-423.

³A. H. Holway and E. G. Boring, Determinants of apparent visual size with distant variant, this JOURNAL, 54, 1941, 21-37.

Procedure. After completing a test of visual acuity, *O* sat in the darkened tunnel with his head braced in a chin-rest by two clamps. A standard set of instructions was read by *E*, indicating that the task was one of size-estimation. It was explained that the time of exposure and the distance to the triangle on *O*'s right would vary from trial to trial. A portion of the instructions read as follows:

I will adjust the size of the triangle on your left at your direction until you believe that its altitude is identical to that of the triangle on your right. Sometimes

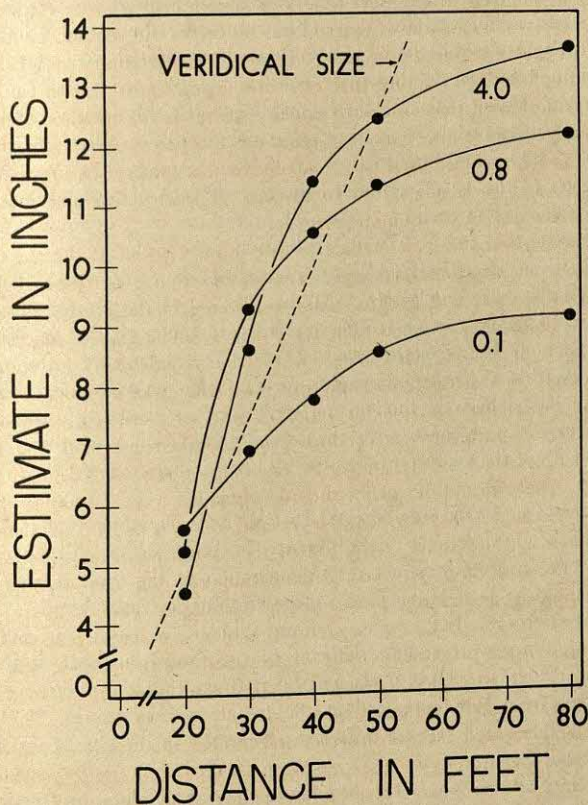


FIG. 1. ESTIMATES OF SIZE OF A STANDARD TRIANGLE AT EACH OF FIVE DISTANCES

All standards subtended 1° of visual angle. Data were obtained under each of three exposure-conditions—0.1, 0.8, and 4.0 sec.

both triangles will be equally distant from you. Often the triangle on the right will be the one farther away. Once in a while it may be the closer one.

Your task is to let me know when you believe the triangles actually have the same height. One way that you could do this would be to imagine that you could pick up the triangle on your left and take it to the place where the other triangle is. If the two would then have the same altitude, you would say that they are

equal. After each exposure you can say "too small" or "too large," until you are satisfied that the two triangles, wherever they are, have the same actual height.

Questions were asked by *E* to ensure that *O* understood his task. About 7 min. were allowed *O* for dark-adaptation.

The *O* was allowed as many viewings as necessary to arrive at a judgment. Usually four or five were sufficient, and these constituted one trial. To minimize anticipatory and perseverative errors, the initial size of the comparison-triangle was so varied that half of the trials were ascending and half were descending. In addition, for both the ascending and the descending approach, the initial altitude of the comparison-triangle was different on each trial. If *O* passed the point of equality he was permitted to reverse direction. The interval between views within a trial depended on the time *O* took to make his judgment. It always exceeded 8 sec. About 1 min. elapsed between trials.

Design. Each of the 12 *O*s estimated the size of each of 5 standards at each of 3 exposures according to each of 2 approach-procedures on 2 occasions. The design, therefore, was $12 \times 5 \times 3 \times 2$ with replication. The order of exposures was as follows: half the 4.0-sec. trials (*A*) were given first; then followed half the 0.8-sec. trials (*B*), all the 0.1-sec. trials (*C*, *C'*), half the 0.8-sec. trials (*B'*), and, finally, half of 4.0-sec. trials (*A'*). This order was used in the hope that it would reduce dependencies among the size-estimates as *O* went from one experimental condition to the next. Within exposure-groups (*A*, *B*, *C*, *C'*, *B'*, *A'*) the approach and the distance to the standard were randomized for each *O*, with the restriction that each standard appeared twice, once with the ascending and once with the descending approach, in each of the six groups.

Results. Analysis of variance indicated that all main variables contributed significantly to the total variance. The difference in the effect of exposure-time, of greatest concern, is quite apparent in the separation of the three curves of Fig. 1. Size-estimates of the standard at 40, 50, and 80 ft. decreased as the duration of exposure was decreased, but estimates of the standard at 20 ft. increased slightly when exposure-times were reduced. Means of the estimates made at the 0.1-sec. exposure for standards at 30, 40, 50, and 80 ft. were found by *t*-test to be significantly different from those obtained at the two longer exposures. The extent to which the situation at 20 ft. was different is reflected in the significant interaction between time and distance.

The curve labeled 4.0 sec. is very similar to the one obtained by Chalmers in an experiment in which *O* had unlimited time for his estimate.⁴ Apparently a 4-sec. exposure permits as effective an estimate of size as does unlimited time. In the present experiment, as time of exposure was reduced, the slope of the curves lessened, indicating that some cues necessary for veridical size-estimation were restricted. Size-constancy was not maintained when accommodation, convergence, and retinal disparity were the only cues to distance *unless* a critical time of exposure was afforded. The minimal interval appears to lie between 0.1 and 0.8 sec. for experimental conditions similar to those reported here.

⁴ Chalmers, *op. cit.*, 415-423.

Leibowitz, Chinetti, and Sidowski provided a cue to distance in the form of a fixation-point and failed to eliminate perspective and interposition as additional cues to distance. It is probable that the differences between their procedure and the present one account for the discrepancy in results. It is interesting to compare the curves based upon their data with those presented here in Fig. 1.⁵ One finds a strikingly similar loss of constancy as exposure-time is reduced. Their curves do, however, suggest a less severe loss.

The data for size-estimation obtained with the 0.1-sec. exposure in the present study indicate clearly that these estimates approach a horizontal asymptote well above the one which would be expected if visual angle were the sole determinant of size. This unexpected regression could be accounted for by a serial effect. It is likely that *O* believed that the standards during the 0.1-sec. exposures continued to be set over the range of distances previously ascertained with the longer exposures. Uncertainty of *O*'s judgment under the briefest exposure would lead him to regress toward the mean of previous estimates.

SUMMARY

Twelve *Os* made estimates of size of triangles at 20, 30, 40, 50 and 80 ft. while limited to accommodation, convergence, and binocular retinal disparity. The role of time in size-constancy was investigated by obtaining these size-estimates at exposures of 0.1, 0.8, and 4.0 sec. The results indicate that an exposure as brief as 0.1 sec. adversely affects size-constancy when only the cues to distance mentioned above are operative.

⁵ Leibowitz, Chinetti, and Sidowski, *op. cit.*, 668.

THE FREQUENCY PROCEDURE FOR ANALYZING THE DATA OBTAINED BY THE METHOD OF LIMITS

By THOMAS MULHOLLAND, Clark University

This paper describes a procedure for analyzing the data obtained by the method of limits which greatly increases, the author believes, the usefulness of the method. This procedure is essentially a new application of an old idea—counting the relative frequency of response for each value of the stimulus. In the interest of definiteness and concreteness of exposition the results of the frequency procedure and the traditional procedure are set side by side. Illustrative data are drawn from Guilford as is the traditional procedure of analyzing them.¹

(1) *Absolute limen.* The data from which the absolute limen is to be calculated by the two procedures are presented in Table I. The signs ↓ and ↑ refer to descend-

TABLE I
DATA OBTAINED BY THE METHOD OF LIMITS ('+'=TONE; '-'=NO TONE)
Series

R (c.p.s.)	1 ↓	2 ↑	3 ↓	4 ↑	5 ↓	6 ↑	7 ↓	8 ↑	9 ↓	10 ↑	11 ↓	12 ↑	13 ↓	14 ↑	15 ↓	16 ↑	17 ↓	18 ↑	19 ↓	20 ↑
14	+		+		+		+		+			+		+		+		+		+
13	+	+	+	+	+		+	+	+	+		+		+		+	+	+	+	+
12	-	-	+	-	-		+	+	+	-		-	+	+	+	+	+	+	+	+
11		-	-	-		+	-	-	-	-	+		-	+	-	+	-	+	-	-
10																				

TABLE II
FREQUENCY TABLE OF JUDGMENTS IN TABLE I

R (c.p.s.)	Frequency		Proportions	
	+	-	+	-
14	20	0	1.00	.00
13	19	1	.95	.05
12	12	8	.60	.40
11	5	15	.25	.75
10	0	20	.00	1.00

ing and ascending series of stimuli, respectively. The stimulus-values (*R*) are cycles per second (c.p.s.) of a lamella—a blade or tongue of steel. The upper and lower *R*-values in Guilford's table have been omitted; only the middle values, those relevant to the present exposition, are given.

By the traditional treatment, the *limit* of each series, *i.e.* the point at which *O* changed his response from '+' to '-' or from '-' to '+', is noted and separately recorded. In these data there are 20 of these points. The mean of these points is regarded as the limen (*RL*) and their *SD* as the measure of the limen's variability. The disadvantage of this treatment is that it tells little about the patterning of the judgments relative to the stimulus-scale.

* Received for publication February 24, 1956.

¹ J. P. Guilford, *Psychometric Methods*, 1955, 101-117.

The procedure advocated here provides this additional information without losing the information obtained by the traditional method. In this treatment it is assumed that if any series were continued beyond the transition point, there would be no reversals, or, that the number of reversals would be small and randomly distributed.

The procedure is as follows: Take any *R*-value in Table I and proceed horizontally along a line which passes through the 20 series. Along this line count the number of '+' and '-' entries. Where there are blanks due to the termination of the

TABLE III

MEANS AND SD OBTAINED BY THE TWO PROCEDURES FROM DATA GIVEN IN TABLE I

Procedures	Mean	SD
Frequency	11.7	0.87
Traditional	11.7	0.84*

*Sheppard's correction applied

TABLE IV

JUDGMENTS OF DIFFERENCE IN BRIGHTNESS BY THE METHOD OF LIMITS

R (degrees white)	VO					VI					VI					VO				
	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑
196°	+		+		+				+		+		+		+		+		+	
192°	+	+	+		+				+		+		+		+		+		+	
188°	+		+		+				+		+		+		+		+		+	
184°						+			+		+		+		+				+	
180°							+													
176°																				
172°																				
168°																				
164°																				

TABLE V

FREQUENCY TABLE OF JUDGMENTS IN TABLE IV

R (degrees white)	Frequency			Proportion		
	+	=	-	+	=	-
196°	24	0	0	1.00	.00	.00
192°	22	2	0	.92	.08	.00
188°	15	9	0	.62	.37	.00
184°	7	17	0	.29	.71	.00
180°	1	19	4	.04	.79	.17
176°	0	13	11	.00	.54	.46
172°	0	6	18	.00	.25	.75
168°	0	1	23	.00	.04	.96
164°	0	0	24	.00	.00	1.00

series, assume that the judgment is the same one as the last given in that series. For example, in Table I, at *R*-value 14 there are 20 '+' and 0 '-'; at *R*-value 12 there are 12 '+' and 8 '-', and so forth. If this procedure is repeated for each *R*-value a frequency table like that presented in Table II is obtained. The frequencies of response are also expressed as proportions of 20.

The mean and *SD* of these functions may be conveniently estimated by Woodworth's summation method.² The values obtained are compared with those given

² R. S. Woodworth, *Experimental Psychology*, 1938, 402-403.

by Guilford in Table III. The agreement is very good. The frequency procedure provides the same statistics as the usual treatment and in addition provides other, complementary information.

(2) *Difference limen*. The data from which the difference limen is to be calculated are presented in Table IV. These data are also from Guilford.³ The letters VO and VI represent two different locations of the standard. The signs ↓ and ↑ refer, as before, to descending and ascending series, respectively. The *R*-values shown represent degrees of white on the variable. *O*'s task was to judge the brightness of the variable relative to that of the standard. A judgment of 'less than' is indicated by '—'; 'equal' by '='; and 'greater than' by '+'.³

By selecting any *R*-value and proceeding along a horizontal line which passes through the 24 series, it is possible to count the number of judgments of each

TABLE VI
MEANS AND SD FROM DATA IN TABLE V.

Responses	Mean	SD
+	186.52	4.08
=	178.88	6.04
—	175.36	5.24

TABLE VII
PSE AND DL COMPUTED BY THE FREQUENCY AND THE TRADITIONAL PROCEDURES

	Frequency	Traditional
PSE	$(\bar{X}_+ + \bar{X}_-)/2 = 180.94$	180.92
DL	$(\bar{X}_+ - \bar{X}_-)/2 = 5.58$	5.58

kind. Where there are blanks due to termination of the series, assume that the judgment is the same as the last one given in that series. It is assumed that if the series were continued beyond the second transition point, there would be no reversals, or, the number of reversals would be small and randomly distributed.

If this procedure is followed for each *R*-value given in Table IV, a frequency table like Table V results. The frequencies are also presented as proportions of 24. The means and SDs of these functions were obtained by Woodworth's summation method. These statistics are presented in Table VI. Using these statistics a *PSE* and a *DL* can be computed. These are presented in Table VII. The values obtained by the traditional method as given in Guilford are presented for comparison. Again, the agreement is excellent.

In conclusion it is clear that the frequency method of analyzing the data obtained by the method of limits is more informative than the traditional method, and, furthermore, that it renders the method of limits the most economical of time of all the psychophysical methods which yield frequency functions of responses. Also, when these frequency functions are used as dependent variables, a more sensitive index of the effect of a given independent variable is obtained, because the whole sample of *O*'s responses is utilized. It might prove interesting to recompare the method of limits, analyzed by the frequency method, with the constant methods, analyzed in the usual way to see if identical operations on the two sets of frequency functions yield identical statistics.

³ Guilford, *op. cit.*, 107.

INTERSENSORY TRANSFER OF VERBAL MATERIAL

By STUART L. WEISSMAN and WALTER H. CROCKETT, Kansas State College

In defending the concept of 'perceptual defense' McGinnies has argued that the relatively high recognition-threshold of taboo words cannot be completely explained in terms of infrequent contact with such words.¹ While he recognizes that taboo words are infrequently seen on the printed page, he says "the critical words are quite common in conversational usage despite their infrequent appearance in print."² To conclude, however, that conversational usage would affect visual

TABLE I

LISTS OF NONSENSE WORDS MATCHED FOR STRUCTURAL PROPERTIES AND LIST OF THE
 PAIRED ASSOCIATES ASSIGNED THEM FOR LEARNING

List A	List B	Associate
lavic	trune	red
dipon	bugew	yellow
jesto	gozie	blue
samig	zonep	orange
acond	encul	white
wovep	ucrag	green
opant	ugenf	purple
harst	kuvel	pink
rofan	vetum	brown
paldo	guiba	black

thresholds requires the assumption that auditory experience will generalize to visual perception. This assumption is specifically rejected by Osgood who says "frequency of conversational usage would not affect recognizability of words as visual stimuli."³

The problem of generalization across sensory modalities has apparently received little experimental attention. Recently Gaydos has shown that Ss who received tactual experience with masonite forms while blindfolded exhibited positive transfer to visual discrimination of the same forms.⁴ The hypothesis underlying the present study was that a similar transfer would occur from auditory experience to visual perception.

Method. Experimental materials were 20 nonsense words of 5 letters each, divided into two lists of 10. The two lists were constructed in such a way that each word in one list was balanced by a word with similar structural properties in the other list. Table I presents the two lists and the associate assigned to each pair. S

* Received for publication July 18, 1956.

¹ Elliott McGinnies, Discussion of Howes' and Solomon's note on "Emotionality and perceptual defense," *Psychol. Rev.*, 57, 1950, 235-240.

² *Ibid.*, 250.

³ C. E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 293.

⁴ H. F. Gaydos, Intersensory transfer in the discrimination of form, this JOURNAL, 1956, 59, 107-110.

learned to respond with the appropriate associates to the oral presentation of the words in one list, and was subsequently tested for speed of recognition of both lists.

Thirty college students, 16 men and 14 women, enrolled in classes in general psychology served as Ss. The first 15 Ss (8 men and 7 women) learned List A; the last 15, List B. In the training trials, *E* read the words aloud to *S*, who was required to respond to each word with the correct paired associate. Order of presentation of words varied for each trial, and a 30-sec. rest was allowed between trials. Learning was continued to a criterion of 20 consecutive correct trials. After achieving the criterion of learning, *S* was dismissed and asked to return the next day.

The transfer test was given approximately 24 hr. after training. *S* was seated next to the projector, approximately 6 ft. from the screen, in a dimly-lighted experi-

TABLE II

NUMBER OF TIMES THE TRAINING AND CONTROL WORDS WERE RECOGNIZED AT SHORTER EXPOSURES

Group	Control word	No difference	Experimental word	Total
A	31	39	80	150
B	36	54	60	150
Total	67	93	140	300

Chi-square: Group A, 15.36, $p < 0.001$; Group B, 3.84 $p < 0.05$; Total, 17.28 $p < 0.001$.

mental room. A barrier was placed beside *S* which prevented him from seeing *E* operate the projector.

Each nonsense syllable was centered on a 16 mm. slide and flashed on the screen by a 220-v. slide-projector equipped with a tachistoscopic shutter. The shutter was set at $f\ 14$, and the projector connected to a 110 v. circuit. Brightness at the center of the screen was approximately 1/10 footcandles.

Before the tachistoscopic presentation was initiated, on the second day, the paired associates learning was resumed to a criterion of five consecutive correct trials. When this criterion was met, *S* was shown three practice slides, consisting of five-letter nonsense words not used on the training lists. Each practice slide was exposed at intervals of 10, 20, 40, and 10 m.sec., respectively. When the practice trials were completed, *S* was informed that practice was over and was encouraged to do his best on the test words.

The slides containing both experimental and control words were numbered 1-20 and presented in random order. The list of 20 words was presented twice at each shutter speed, beginning with 10 m.sec and moving to 20, 40, 60, 200, 500, and 1000 m.sec. When a word was correctly identified at two consecutive exposures it was eliminated from the series. The shutter speed of the second correct recognition was taken as the recognition-threshold.

Results. The effect of training was tested by determining the frequency with which each word on which *S* received training was recognized at a speed of exposure faster than, equal to, or slower than its structural counter-part. Table II presents this comparison for each experimental group and for both groups combined.

It is evident that the Ss in Group A showed more transfer of auditory training than did those in Group B. This probably resulted from the fact that the words in List A were somewhat easier to recognize than those in List B. Even for Ss in Group B, however, the training syllables were recognized at a shorter exposure significantly more often than the control syllables, and the results for the two experimental groups combined are clearly significant.

Discussion. These results indicate that transfer does occur from auditory training to visual discrimination. The processes mediating the transfer are not, however, clear. It is possible, for example, that the Ss gained insight into the nature of the experiment by recognizing certain of the training words early in the tachistoscopic series, and subsequently looked for other words in the training group. Since the order of tachistoscopic presentation of words was not counterbalanced, it was not possible to determine the extent to which the results may be explained by the operation of such a set. An experiment is currently in progress to answer this question.

Even if the results were completely determined by the operation of set, however, they still indicate that Ss with the concerted training in visual recognition of words that is provided in a society such as ours, are able to generalize from auditory experiences to the visual discrimination of words; otherwise, the formation of a perceptual set would be impossible. As McGeoch and Irion say, "modality of stimulation does not wholly determine the subject's apprehension of the material. . . . If unaccustomed to auditory presentation, he may attempt to imagine the material visually or to speak it subvocally or he may make implicit movements of writing or drawing it. The receptor is the starting point of the practiced response, but it is by no means its sole determiner."⁵

⁵ J. A. McGeoch and A. L. Irion, *The Psychology of Learning*, 1952, 481.

MANIFEST ANXIETY, DISCRIMINATION, AND TRANSPOSITION

By A. W. BENDIG and CHARLES J. VAUGHAN, University of Pittsburgh

Stevenson and Iscoe have recently reported the marked superiority of low- (*LA*) as compared with high-anxiety (*HA*) Ss (college students selected on the basis of Taylor's scale) in the discrimination of intermediate size.¹ Their results are in accord with those of Hilgard, Jones, and Kaplan for conditioned discrimination,² but not with those of Grice for discriminative reaction-time (*RT*),³ or with those of Eriksen and Wechsler for the discrimination of size under conditions of induced anxiety⁴—in neither of the latter studies did a significant difference appear. The present experiment represents an attempt to replicate the results of Stevenson and Iscoe, as well as to extend their procedure to the study of transposition, another area in which contradictory findings have been reported. Maltzman, Fox, and Morrisett found a relation between Taylor score and transposition,⁵ while Wesley could find no such relation.⁶

Subjects. Three psychometric tests were administered at the beginning of the semester to 168 Ss enrolled in a one-semester introductory course in psychology. The tests were: (a) a modified 20-item form of the Taylor Manifest Anxiety Scale; (b) a 20-item vocabulary test drawn from the Coöperative Vocabulary Test; and (c) the 28-item Need Achievement Scale of Edwards' *Personal Preference Schedule*.⁷ Raw scores on each of the scales were converted to stanines based upon university or national norms. A total of 40 Ss were selected from the top and bottom three-stanine extremes of the Taylor distribution, with 20 Ss in the *HA* group (raw scores of 9 or more anxious responses) and 20 Ss in the *LA* group (raw scores of 2 or fewer anxious responses). One additional *LA* S was trained on the task, but failed to reach criterion in over 2-hr. of testing and had to be replaced by another. There were 31 men and 9 women, with 15 of the men in the *HA* group. All of the selected Ss were talked to individually in class near the end of the semester and

* Received for publication August 1, 1956.

¹H. W. Stevenson and Ira Iscoe, Anxiety and discriminative learning, this JOURNAL, 69, 1956, 113-114.

²E. R. Hilgard, L. V. Jones, and S. J. Kaplan, Conditioned discrimination as related to anxiety, *J. exp. Psychol.*, 42, 1951, 94-100.

³G. R. Grice, Discrimination reaction time as a function of anxiety and intelligence, *J. abnorm. soc. Psychol.*, 50, 1955, 71-74.

⁴C. W. Eriksen and Henry Wechsler, Some effects of experimentally induced anxiety upon discrimination behavior, *ibid.*, 51, 1955, 458-463.

⁵Irving Maltzman, Jack Fox, and Lloyd Morrisett, Jr., Some effects of manifest anxiety on mental set, *J. exp. Psychol.*, 46, 1953, 50-54.

⁶Elizabeth Wesley, Perseverative behaviors in a concept-formation task as a function of manifest anxiety and rigidity, *J. abnorm. soc. Psychol.*, 48, 1953, 129-134.

⁷A. L. Edwards, *Manual for the Edwards Personal Preference Schedule*, 1954, 1-18.

requested to participate in the experiment, although none of them was informed that his selection was related to his test-performance of three months before.

Apparatus. The apparatus consisted of: (a) four $3 \times 3 \times \frac{3}{4}$ -in. blocks painted flat black, each bearing a white cardboard square; (b) a flat black tray containing three wells for presenting the blocks; (c) a screen lowered between trials to shield E; and (d) a white poker chip which served as the reinforcing stimulus. The areas of the white squares were $\frac{3}{4}$, 1, $1\frac{1}{2}$, and $2\frac{1}{4}$ sq. in., the first three squares being presented on each of the learning trials, and the $2\frac{1}{4}$ -sq. in. card replacing the $\frac{3}{4}$ -sq. in. card on the single transposition trial.

Procedure. Each S was seated at a table opposite E who demonstrated how the blocks could be used to cover the wells on the board. The instructions were the same as reported by Stevenson and Iscoe.⁸ The poker chip was always under the block bearing the middle-size (1-in.) square during the learning trials with a randomized positional sequence being used for each S. Practice was continued to

TABLE I
TRIALS TO CRITERION FOR ANXIETY AND TRANSPOSITION GROUPS

Group	N	Mean trials	Median trials	Mean trials
High anxiety	20	16.1	4.5	0.920
Low anxiety	20	12.3	6.2	0.858
Transposers	17	9.5	3.7	0.718
Nontransposers	23	17.7	7.7	1.016
Total	40	14.2	5.8	0.889

the criterion of five successive correct choices. On the trial following the fifth criterion-trial, the $\frac{3}{4}$ -in. square was replaced by the $2\frac{1}{4}$ -in. square and Ss were classified as transposers (selecting the $1\frac{1}{2}$ -in. square) or as nontransposers (selecting the previously reinforced 1-in. square).

Results. On the transposition trial, 8 of the HA Ss as compared with 9 of the LA Ss transposed, a difference that is obviously not significant. Since the distribution of the trials to criterion (omitting the criterion-trials) was highly skewed in the positive direction, the scores were transformed by adding one to each score and converting it to a common logarithm, and a two-criterion analysis of variance was performed with the Ss' being classified as HA or LA and as transposers or nontransposers. Similar analyses of variance were made for the Vocabulary and Need Achievement stanine-scores. The log-trials scores were adjusted for Vocabulary and Need Achievement separately and also for the combination of the two tests by covariant procedures. None of the four analyses gave evidence of a significant difference in the learning performance of the HA and LA groups; in fact, all of the relevant *F*-ratios were less than 1.0. The difference in rate of learning between the transposers and nontransposers was significant only at the 10% level in all four analyses. It is to be noted that the covariance adjustments had little effect upon the results for the unadjusted log learning scores, an outcome which was reflected in

⁸ Stevenson and Iscoe, *op. cit.*, 113-114.

the intercorrelations between the three measures: log learning trials and Vocabulary, $r = 0.15$; log learning trials and Need Achievement, $r = -0.01$; Vocabulary and Need Achievement, $r = 0.15$.

Rates of learning in the anxiety and transposition groups are given in Table 1.

Discussion. The history of the research with Taylor's scale is dogged by contradictory results. For example, compare the results of Taylor and Spence⁹ with those of Hughes, Sprague, and Bendig¹⁰ on serial learning, or the results of Farber and Spence¹¹ with those of Axelrod, Cowen, and Heilizer on stylus-maze learning.¹² The same phenomenon appears when the results of the present study are compared with those of Stevenson and Iscoe. One difference between the two experiments is in the form of the anxiety scale used. Stevenson and Iscoe used the usual 50-item scale, while we used a 20-item form that had been devised by selecting only those items from the standard scale that two previous item-analyses had shown to be related to clinical manifestations of anxiety.¹³ Our form of the test does, however, correlate highly ($r = 0.93$) with the total scale (of which, of course, it is a part). The general impression one receives from reviewing the results above is that the relationship between the Manifest Anxiety Scale and learning is, like extrasensory perception (ESP), a delicate flower that blooms only in certain environments.

⁹ Janet A. Taylor and K. W. Spence, The relationship of anxiety level to performance in serial learning, *J. exp. Psychol.*, 44, 1952, 61-64; Spence and Taylor, Erratum, *J. exp. Psychol.*, 46, 1953, 462.

¹⁰ J. B. Hughes, J. L. Sprague, and A. W. Bendig, Anxiety level, response alteration, and performance in serial learning, *J. Psychol.*, 38, 1954, 421-426.

¹¹ I. E. Farber and K. W. Spence, Complex learning and conditioning as a function of anxiety, *J. exp. Psychol.*, 45, 1953, 120-125.

¹² H. S. Axelrod, E. L. Cowen, and Fred Heilizer, The correlates of manifest anxiety in stylus maze learning, *ibid.*, 51, 1956, 131-138.

¹³ Bendig, The development of a short form of the Manifest Anxiety Scale, *J. consult. Psychol.*, 1956, 20, 1956, 384.

VERBAL COMPARED WITH MANIPULATIVE SOLUTION OF AN APPARATUS-PROBLEM

By WILBERT S. RAY, Air Force Personnel and Training Research Center

This experiment asks whether Ss may solve problems in a different way if they are forced to talk, and so presumably to think, about them before doing anything else about them. It is related to an experiment by Saltz, Moore, and Hoehn, which suggests that Ss do a better job of finding sources of trouble in electrical equipment if they 'pre-plan' (*i.e.* make a plan for) their work.¹

Subjects. The Ss were 128 basic airmen at Lackland Air Force Base. Their scores on the Armed Forces Qualification Test place them roughly in the upper two thirds of the general population in intelligence. They were assigned to one of two groups in random order.

Task. The equipment employed contains a vertical panel carrying a light which can be turned off by throwing simultaneously two of the seven toggle-switches on an inclined response-panel. Each S was requested to turn off the light. The Ss in the experimental group were then told, "But, before you actually touch any of the switches I want you to tell me what you are going to do." The latter statement usually had to be supplemented by such questions as "Exactly which one?" or "And what would you do next." Only after the verbal statement did these Ss begin to manipulate the switches. The Ss in the control group did not have to make any verbal statements before starting to work on the problem.

Manual behavior was recorded mechanically and verbal behavior by the E. The experimental Ss were encouraged to produce as many verbal trials as possible, with complete silence for 60 sec. accepted as the criterion for terminating this phase of the work and proceeding to the manipulative phase. No information about the adequacy of the verbal formulations was provided by E.

Results. Extinguishing the light required the simultaneous use of two switches. Once the multiple-switch hypothesis had appeared in S's behavior, there was no rational basis for preferring any one pair of switches to any other pair, and the actual number of trials between the appearance of the hypothesis and the desired outcome was purely a matter of chance. The mere appearance of the hypothesis, therefore, was adopted as the criterion of success, although S was not so informed. All the Ss in both groups were successful in the manipulative solving, but the level of success in the verbal work of the experimental group was only 73%.

The median number of trials preceding success was 7.78 for the verbal work of the experimental group, 7.26 for the manipulative work of the same group, and 8.71 for the (manipulative) work of the control group. Medians are used because the data are skewed, and because they permit the inclusion of the records of the 17 Ss unsuccessful in the verbal phase who were given a trial-score of infinity. The difference in number of verbal and manipulative trials to solution in the

* Received for publication July 19, 1956. This research was carried out at the Air Force Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas, in support of Project 7709.

¹ Eli Saltz, J. V. Moore, and A. J. Hoehn, Test of a theoretical formulation relating a preplanning technique to factors in successful problem solving, *TARL Laboratory Note* 55-5, 1955, (unpublished).

experimental group was significant at the 1% level (sign test). Chi-square tests of the other two pairs of medians showed the differences between sets of data to be significant at or beyond the 5% level. (All differences reported below were found to be significant at the 5% level or better on the basis of chi-square tests.) Success was achieved with a smaller number of trials in manipulative than in antecedent verbal work, and in a smaller number of trials in manipulative work preceded by verbal work than in manipulative work not so preceded.

The hypothesis that a single switch will extinguish the light is the simplest hypothesis which *S* could use, and it demands the smallest number of trials (seven) for complete exploration. If we may define the most systematic attack as the use of seven single switches before starting to work on a more complex hypothesis, how systematic was the work under the different conditions? In the verbal work 63% of the *Ss* were systematic; in the manipulative work 36% of the experimental *Ss* and 20% of the control *Ss* were systematic. That is, manipulative work was less systematic than antecedent verbal work, and the verbal work made *Ss* more systematic in later manipulative work than they would otherwise have been.

The *Ss* who were unsuccessful in the verbal work had a smaller total number of verbal trials than did the successful verbal solvers, the medians being 11.5 and 19.5 trials, respectively. It will be remembered that these *Ss* had no information about whether or not they were successful. The median number of trials per *S* for the 64 *Ss* in the experimental group during the verbal work was 17.0, and the median of the trials made by the *Ss* of the control group was 25.3. Unsuccessful verbal solvers made smaller number of trials than did successful *Ss*; and verbal solvers made fewer trials than did manipulative solvers.

Discussion. Verbal behavior was more systematic than the manipulative behavior of either group. It may be that in verbal trials, which took more time than did manipulative trials, *S* was more aware of the fact that he had tried a given switch and therefore less willing to try it again. In any event, it is not surprising that the systematic behavior persisted in the later manipulative behavior of the same *Ss*.

In two of the three comparisons made above, the difference in amount of systematic behavior was reflected in an inverse difference in number of trials preceding success. More systematic behavior should be accompanied by a smaller number of trials, since our definition of systematic behavior means no repetition of error on the single switches. In the case of the third comparison (the verbal and manipulative behavior of the experimental group) the decrease in number of trials in the manipulative behavior is presumably due to previous practice with the task and to the corresponding increase in the amount of time spent on it.

The smaller total number of trials in the verbal behavior and the (resulting?) smaller degree of success may very well be due to the lack of knowledge of results here, and if so these are features of the particular procedures used rather than features of verbal behavior in general.

Summary. A group of 64 basic airmen solved an equipment-problem which resembles a search for a malfunction. Another group solved the problem, but only after its members told *E* what they would do before they actually touched the apparatus. The *Ss* of the latter group made fewer repetitive errors and required fewer trials than did the purely manipulative solvers.

APPARATUS

APPARATUS FOR THE STUDY OF VISUAL TRANSLATORY MOTION

By OLIN W. SMITH and JAMES J. GIBSON, Cornell University

A psychophysics of the visual perception of motion requires apparatus for the control and systematic variation of the optical stimulus. The effective stimulus is some change in the pattern of light rays entering an eye. While this is ordinarily produced by physical motion in the environment (or physical movement of the observer) it need not be so; a change of this sort can also be produced by various devices such as shadow-casters, motion pictures, and a moving surface behind an aperture or slot.¹

Perhaps the simplest type of optical movement is translation, which can be defined as linear displacement of a figure whose parts all have the same amount and direction of displacement. It is produced, for example, by a solid object or surface moving in a direction perpendicular to the line of sight. Translation, like rotation, is a geometrically 'rigid' motion. It can be conveniently presented to an *O* by arranging for an object or a surface to move behind an aperture or window in a screen which is viewed in the frontal plane. A pair of kymograph drums (standard and variable) carrying an endless belt on which patterns are inked has been frequently used.² The limitations of this set-up, however, are obvious. (1) The moving fields are small; (2) the speed of movement is only roughly controlled; and (3) the range of stimulus-conditions is narrow. There is need for an apparatus that is accurate; that permits of a wide range in the size of the stimulus-fields and in the speed of their movement; and that is conveniently used. With these points in mind the following specifications were drawn up: (a) The dimensions of the moving fields should be so large that a stimulus-field of wide visual angle may be used when desired. An

* The construction of this apparatus was supported in part by contract AF33(038)-22804 between Cornell University and the Audio-Visual Research Division, Human Resources Research Laboratory, Bolling Air Force Base, Washington 25, D.C., and contract Nonr 401(14) between Cornell University and the Office of Naval Research, Department of Navy, Washington 25, D.C.

¹ J. J. Gibson, The visual perception of objective motion and subjective movement, *Psychol. Rev.*, 61, 1954, 304-314.

² J. F. Brown, The visual perception of velocity, *Psychol. Forsch.*, 14, 1931, 199-232.

area 3×4 ft. was decided upon. From this maximum, the area may easily be reduced by varying the size of windows in a screen. (b) The range in the speed of movement should be large: from 0-1000 ft. per min.; and the direction of movement should be reversible at the will of *E*. (c) The

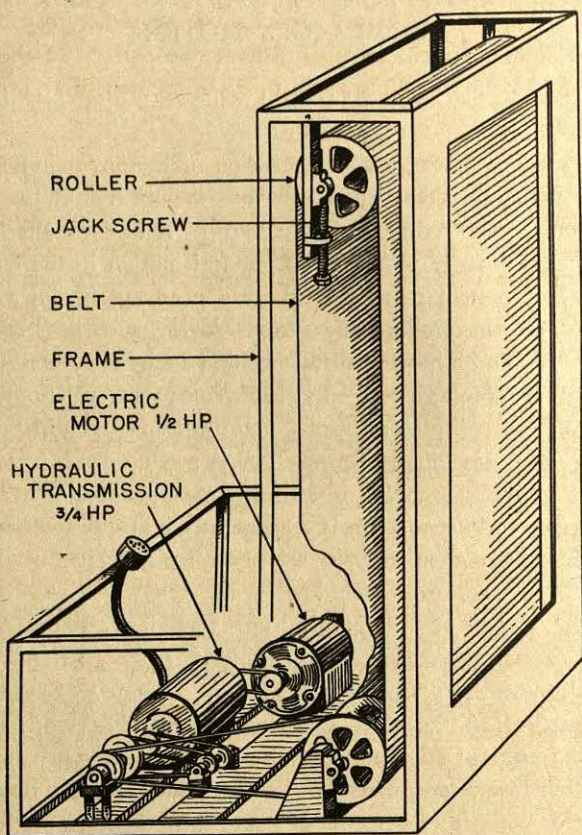


FIG. 1. SCHEMATIC LINE DRAWING OF THE APPARATUS FOR TRANSLATORY MOTION

velocity of the surface should be variable by continuous rather than discontinuous steps, and the variations should be easily made. (d) *O* should be provided with a remote control of the variable apparatus and *E* should have remote control of *both* the standard and the variable. The true velocity of both should be easily determinable by *E*. (e) If the driving mechanism could not be made noiseless (as proved to be the case) the sound of the mechanism should not be a simple function of the velocity of

the belt. (f) The distance of the moving surface from O should be easily variable, and the plane of the moving surface should be variable with respect to tilt (rotation) and slant (inclination to the line of sight). (g) The moving fields (belts) should be capable of being changed easily.

The apparatus described here was constructed to meet these specifications (see Fig. 1). The two pieces are identical, hence either may be used as the standard or the variable. Each consists of two parts: (1) the belt-and-window, and (2) the mechanism which drives the belt. Both are mounted in a single frame of welded angle iron.

The moving surface. Each belt was made of three-ply cotton cord, rubber backed on both sides, diagonally spliced so as to rotate smoothly.^a Each is 115 in. long by 36 in. wide; when mounted, the largest observable area is a flat surface 36×48 in. Two axles rotate in friction, self-positioning bearings fastened to the frame. They are parallel and in the same plane. Each bearing of the top axle is adjustable for elevation by jack screw mounts which control tension on the belts. Upon each axle seven $5\frac{1}{2}$ -in. pulleys are equally spaced between two flanking 6-in. pulleys. These nine pulleys on each axle provide the support over which rotates the belt. Around the inside edges of each belt two $\frac{5}{8}$ -in. V-belts are riveted. These V-belts ride in each of the two exterior 6-in. pulleys which are mounted at either end of each axle, and serve to prevent creeping and to hold the belt in place when the apparatus is at varying degrees of slant or tilt. The bottom axle (the driving axle) is non-adjustable. It is driven by a V-belt over a tenth pulley.

Metal grooves on the front of the frame can hold a plate glass window which completely covers the available observational surface. The window is about 6 in. in front of the belt, and can support fixation-points or masks for any position of the moving field. Various designs may be displayed on the moving surface by numerous methods. Wall paper, for instance, can be employed. The front of the apparatus surrounding the window is shielded by a black metal covering.

The drive. The principal components of the drive assembly are a $\frac{1}{2}$ HP electric motor coupled to a Vickers (TR3) $\frac{3}{4}$ HP variable speed hydraulic transmission. The output shaft is connected to the bottom axle by V-belts and pulleys. Several different types of remote drive mechanisms were possible for controlling the speed of the transmission. It was found simplest to attach selsyn motors wired through a switching panel. The maximal velocity of the belt is about 2000 ft. per min.

The transmission of each piece has a speed-control post which rotates through a zero point of no output. Twenty-four turns of the speed control post reverse the direction of rotation from a maximum in one direction to maximum in the other direction. The maximal belt-speed is alterable, however, by installing other pulleys with different pulley ratios. The amount of change in velocity per turn of the control post is therefore a function of the installed pulley ratio. Either E or O can operate the speed control post, at the option of E , by means of selsyn units, one of which rotates the speed control post of the transmission. The velocity

^a Manufactured according to our specifications by the Goodyear Co. of Akron, Ohio.

of the belt is easily measured by a tachometer held by hand or permanently mounted.

The apparatus is mounted on casters in the vertical position; it is cradled on a caster mounted platform when in a horizontal position; and it can be propped by jacks and blocks at other positions of slant or tilt.

Use of the apparatus. The apparatus is a flexible device which makes possible the study of many problems concerned with the perception of visual motion. When a pair of identical windows is presented to an *O*, the method of adjustment or reproduction can be used. Some of the problems to be investigated are as follows.

(1) The effects of any of the following on the accuracy (variability) of reproduction of velocity: (a) the separation of the windows in *O*'s visual field; (b) the velocity of the standard (Weber's law); (c) the direction of the moving fields (right, left, up, or down); (d) the mode of observation employed in comparison (stable or moving fixation of the eyes); (e) the part of the retina stimulated (central or peripheral) with stable fixation of the eyes; (f) the training and age of *O*; (g) the size and kind of pattern or texture imposed on the moving belts.

(2) The investigation of the constant errors when *O* matches the variable and standard fields. (a) Is vertical motion phenomenally slower or faster than horizontal? (b) Is motion perceived with fixation of the field phenomenally faster than with pursuit eye-movements (the Aubert-Fleischel paradox)? (c) Is motion perceived in a smaller window phenomenally faster than in a larger window? (d) Is motion perceived in a distant window phenomenally the same as in a near window of the same or different size?

(3) The effects of practice on the discrimination of velocity and absolute judgments of velocity. (a) How is the upper threshold of visual speed (phenomena of 'streaming,' blurring, flicker and flicker fusion, etc.) affected? (b) Is the recognition of forms imprinted on the belt dependent on the speed of movement? (c) Does the apparent direction of motion depend on the shape and orientation of the window framing it?

CONTINUOUS RECORDING OF THE FLUID-INTAKE OF SMALL ANIMALS

By PAUL THOMAS YOUNG, University of Illinois

Psychologists and physiologists concerned with the study of thirst or the intake of various solutions of sugar, salt, vitamins, or minerals, require a sensitive apparatus that can register not only the times at which an animal drinks but also the quantities of fluid ingested during the experimental periods. The device described here has been used in our laboratory for more than a year and it has been found to be highly sensitive and satisfactory in every respect.

Other techniques for recording or controlling the intake of animals have been described. Gregersen pictured a "potometer" for recording the drinking behavior of dogs.¹ His instrument consisted of a water tank suspended by a coiled spring with marker attached to write on a kymograph paper; this tank was connected by a rubber tube to a drinking cup with an arrangement for maintaining an approximately constant level of liquid in the cup. This device, he reports, was not very sensitive because, "Removal of less than 25 cc. of water from the drinking pan is not recorded on the drum."² A more sensitive apparatus was described by Young and Richey for continuous recording of the fluid-intake of rats.³ With this apparatus the writing pen dropped in a step-wise fashion rather than in a smooth curve; from 2-3 cc. of fluid had to be removed from the cup before the pen moved.

A very different technique requires the animal to work by pressing a bar for a small quantity of food (liquid or solid). Skinner described a valved siphon arrangement for delivering an approximately constant quantity of water to a watch crystal every time the rat pressed a bar.⁴ Guttman and Estes devised a dipper mechanism to bring up to the rat a constant quantity of liquid or solid food every time bar pressing closed the circuit.⁵ These instruments dole out unit quantities of reward and automatically register the frequency of bar pressing.

Another method of studying drinking behavior depends upon the closing of an electronic circuit through the animal's body. An electric current, of subthreshold intensity, is passed from the floor of the cage through the tongue of the test-fluid. Every tongue contact with the fluid closes the circuit. Such an apparatus was devised by Kappauf for recording tongue responses of rats, dogs, and cats and

¹ M. I. Gregersen, Studies on the regulation of water intake: II, Conditions affecting the daily water intake of dogs as registered continuously by a potometer, *Amer. J. Physiol.*, 102, 1932, 344-349.

² *Ibid.*, 344.

³ P. T. Young and H. W. Richey, Diurnal drinking patterns in the rat, *J. comp. physiol. Psychol.*, 45, 1952, 80-89.

⁴ B. F. Skinner, Thirst as an arbitrary drive, *J. gen. Psychol.*, 15, 1936, 205-210.

⁵ Norman Guttman and W. K. Estes, A modified apparatus for the study of operant behavior in the rat, *J. gen. Psychol.*, 41, 1949, 297-301.

for studying species differences in the lapping of water.⁶ The same principle was utilized in a "drinkometer" described by Hill and Stellar.⁷ From the frequency of tongue contacts they were able to plot intake curves for rats in a free-drinking situation. They calculated that 0.004-0.005 ml. of fluid was ingested per tongue lap.

The apparatus described here resembles that of Young and Richey more closely than any of the other types. It, in fact, makes use of the same kymograph and recording pens but the drinking cup has been eliminated and a more usual drinking spout with nipple has been substituted. Although tongue contacts are not recorded electronically, this feature could be added if desired.

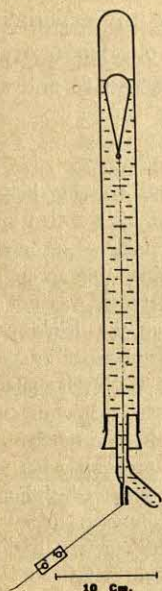


FIG. 1. CROSS SECTION OF
A RECORDING UNIT

The plan of the intake recorder can be easily understood from Fig. 1. The essential part of the unit is the drinking spout modified by fusing a capillary tube to it in such a position that a filament may be drawn through the capillary from liquid into the air. This boot-shaped drinking spout is clamped firmly to the rat's cage with the nipple inside at a height convenient for the rat (5-8 cm. above the floor). The large tube containing the test-fluid is supported rather loosely near the top.

Measurements and specifications are as follows: The inverted test-tube is 40 cm. in length and made of pyrex glass, 32 mm. in outside diameter (OD). The spout is made of pyrex tubing, 10 mm. in outside diameter, with a 3 mm. aperture at the nipple. Fused to the spout, as shown in Fig. 1, is a capillary tube with a bore of 0.5 mm. The capillary is 6 mm. in outside diameter and 10 mm. in length from seal to the end. The opening of the capillary tube is level with the upper edge of the aperture at the end of the spout.

An important feature of the apparatus is the use of a plastic monofilament that has a negligible rate of absorption. The reason this is important is that ordinary threads, including nylon, are slightly longer, under a

⁶ In a footnote to his chapter on "Discriminative behavior in animals" Dr. K. U. Smith gives credit for the original design of this type of apparatus to Dr. William E. Kappauf. See C. P. Stone (ed.), *Comparative Psychology* (3rd. ed.), 1951, 334-335.

⁷ J. H. Hill and E. Stellar, An electronic drinkometer, *Science*, 114, 1951, 43-44.

constant load, when wet than when dry. Since the filament passes steadily from water to air, it is important that the length remain the same when wet and dry.⁸ The filament can be pulled down vertically to a recording device but we have drawn it at an angle thus making a single friction-point at the base of the capillary tube. The free end of the filament is attached to a celluloid adjustor which, in turn, is connected to an ordinary thread that leads to the recording pen, as described elsewhere.⁹

We are presently using glass floats but formerly used cork floats. With cork floats there was occasionally some irregularity in the action caused by air bubbles adhering to a cork. The float shown in the illustration has an outside diameter of 22 mm. at the upper end. It has an over-all length of

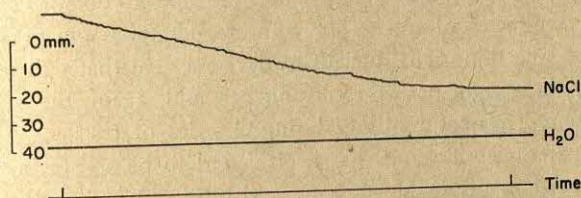


FIG. 2. CURVES OBTAINED IN A 10-MIN. PREFERENCE TEST WITH TWO RECORDING UNITS

One unit contains a 0.9% solution of NaCl and the other distilled water.

63 mm. and is made of thin-walled (about 0.5 mm.) glass. A small copper wire is fused into the glass at the lower end of the float for attachment of the filament.¹⁰

In preparing the apparatus for use we first fill the inverted test-tube with water or a solution. Then the float and stopper are inserted. If there should be air leakage when the tube is turned over, a little vaseline generally stops it. When the apparatus is properly adjusted small bubbles of air rise in the tube as the rat laps the fluid at the nipple. Each air bubble causes the float (and consequently the pen) to settle a bit.

At the present time we are operating six of these self-recording intake units simultaneously and registering six curves on the same kymograph paper. A sample record is reproduced in Fig. 2. This figure shows tracings of a pair of curves obtained during a 10-min. preference-test between two

⁸ The filament was sent to us through the courtesy of E. I. duPont, deNemours and Co., Arlington, New Jersey.

⁹ Young and Richey, *op. cit.*, 80-83.

¹⁰ Glass parts of the apparatus were made by Mr. A. F. Robbins, senior glass-blower at the University of Illinois, whose skill is largely responsible for the successful operation of the equipment.

fluids presented simultaneously. The S was a *nonthirsty* adult rat, well habituated to the testing cage and to the drinking spouts before the record was made. One of the units was filled with a 0.9% NaCl solution and the other with distilled water. At the time of the test the rat was placed in the cage as indicated on the time-line and removed exactly 10 min. later. The curve shows in millimeters the fall of the pen (and float) during the experimental interval. By calibrating the fall of the pen, the fluid-intake may be computed in cubic centimeters. Incidentally, this record shows that a *nonthirsty* rat, when offered a choice between distilled water and 0.9% NaCl solution, spends most of his time drinking the salt solution and little, if any, drinking the water.

I have used an apparatus with smaller test tubes to obtain a larger drop per cubic centimeter of intake but with smaller tubes practical problems appear. For one thing it is difficult to devise a float that will not be submerged by the weight of the recording pen and, again, the smaller tube requires more frequent refilling during a series of tests.

The apparatus as it stands has been found to be satisfactory for continuous recording of intake of water and other fluids. I believe that the boot-shaped spout can be adapted to many uses when continuous records of fluid intake are desired.¹¹

¹¹ Another type of apparatus has been described since completion of the present paper. See M. L. Duffy and G. E. Price, A continuous drinking-recorder for small animals, this JOURNAL, 69, 1956, 664-666.

A CONTINUOUS MILLISECOND CONTROL FOR THE GERBRANDS TACHISTOSCOPE

By ROSS STAGNER, H. C. ROBERTS, M. L. GARDNER, and
ARCHER L. MICHAEL, University of Illinois

The apparatus here described has been used with the well-known Gerbrands tachistoscope¹ to produce continuously variable exposure-intervals from 6-580 m. sec. Using the light-sources built into the Gerbrands, we first devised a control involving electronic determination of time-intervals and vacuum tubes as switching devices. This form of the apparatus proved unreliable in operation because the fluorescent tubes require the steep wave-front (usually several times repeated because of contact-bounce) which is produced by a mechanical switch. In the final form of the apparatus, time-intervals still are determined electronically, but the switching is done with a small and reliable high-speed relay (Sigma Type 4F) which has a response-time considerably less than 1 m. sec.

The time-intervals are generated by the one-shot multivibrator, shown with its power-supply in Fig. 1. The unit employs a 6SN7 electron tube, containing two triodes in a single envelope. In its normal or quiescent condition, Triode V_1 permits current to flow through the Sigma 4F relay, connecting Contacts 2 and 3, thus closing the electrical circuit to illuminate the steady field. Operation of Sw_1 places a very short pulse of positive polarity on the control-grid of Triode V_2 , causing it to conduct current. This sudden rush of current through V_2 and its 20,000 Ω resistor produces a large negative pulse at the plate-terminal of V_2 ; this pulse is applied through a capacitor to the control-grid of V_1 , causing this tube to stop conducting current and the relay to operate. In this portion of the operating cycle, V_2 rather than V_1 is conducting current, the relay is opened rather than closed, and the stimulus-field rather than the preexposure-field in the tachistoscope is illuminated.

The duration of the timing cycle is controlled by the RC -constant of the circuit involving Sw_2 and Sw_3 . These two multi-point switches permit the selection of five different capacitors ranging from 0.07-1.0 mfd., and resistances which vary from 0 to 600,000 Ω , in steps of 120,000 Ω (interpolated values are obtained through the variable resistance shown).

¹ Ralph Gerbrands, Arlington, Massachusetts.

The operating cycle ends when the negative pulse applied to the grid of V_1 has been dissipated through the resistance at Sw_3 , and at this time current flows again through V_1 , actuating the relay and transferring the illumination to the preexposure field. The circuit is quite accurate and reliable, and not much affected by the characteristics of the electron tube. Calibration data are plotted in Fig. 2.

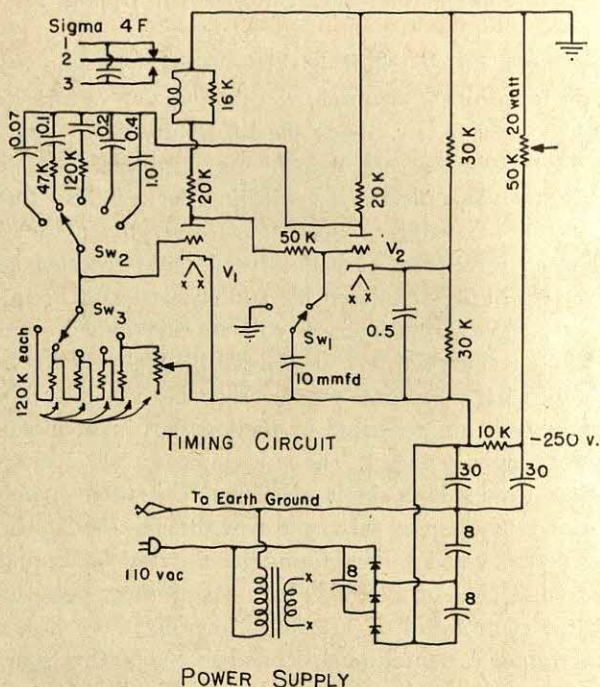


FIG. 1. WIRING DIAGRAM FOR ILLINOIS TIMER

Sigma 4F, relay; V_1 - V_2 , triodes of 6SN7 tube; 0.07, 0.1, 0.2, etc., fixed capacitors; 120K, series of fixed resistances plus variable resistance; Sw_2 , selector for capacitors; Sw_3 , selector for resistances (there is also a control knob for the variable resistance); Sw_1 , initiating switch; 1, 2, 3, leads to lighting circuit.

The circuit of Fig. 1 shows a power-supply that is somewhat unusual in that its positive terminal instead of its negative is connected to ground. This connection was necessary to permit the operation of the purely electronic form of the timer. Since the unit as used did not employ this added circuitry, it has not been included in the drawing. It is for this reason that one of the power leads is marked "to earth ground"; this insures safety in operation under all conditions.

Calibration was accomplished by the use of a photoelectric cell (General Electric catalogue number 88X565). The impulse from this cell was connected to the Y-input of a DuMont 304A oscilloscope with controls set at a nominal value of D.C. gain for the Y-input amplifier. The sweep used was a driven sweep, *i.e.* initiated by each Y-axis pulse. The sweep-rate was set at a very low point and the X-amplifier adjusted to give a complete picture of the various intervals. It was thus possible to plot the fluorescent tube's build-up, full-on time, and decay-interval.

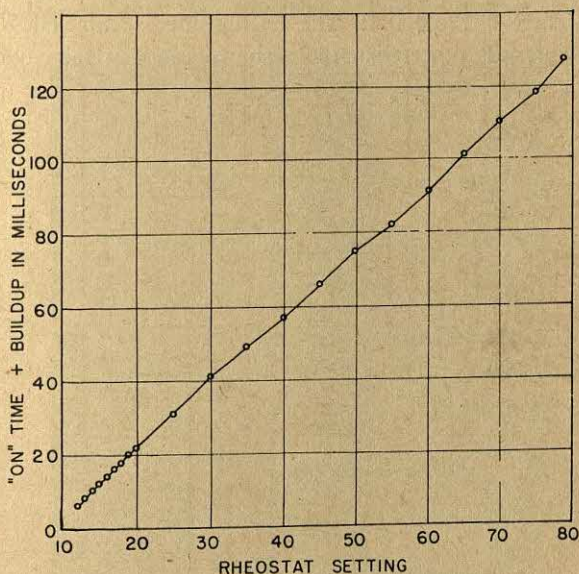


FIG. 2. RELATION OF RESISTANCE TO ILLUMINATION-TIME

"Rheostat setting" is in arbitrary units varying the resistance at S_{w3} with no fixed resistors in circuit; illumination-time is measured in milliseconds.

To determine time-intervals, an audio-frequency square wave was applied to the input of the oscilloscope. A Heath Kit Audio Generator (Model A6-7) was used. A frequency of 1000 $\sim$ on the audio would cause dots or streaks, each equalling 1 m. sec., on the screen; a frequency of 250 $\sim$ would give dots equalling 4 m. sec., and so forth. A Grass photographic recorder was used to make permanent records of the traces.

The full build-up time for the light-sources in both of the Gerbrands instruments tested was 11 m. sec. and highly constant. When the Illinois timer was set for shorter periods, the lights did not reach full-on. The

slow decay of these light-sources also raises a serious problem as to the actual 'exposure' of the stimulus-field, *i.e.* the point at which the obscuring effect of the illumination in the preëxposure field is overcome by the light for the stimulus-material. Not less than 12 m. sec. were required for the first 50% of decay of one light-source, and not less than 17 m. sec. for the other one tested. (The variability of decay-time also introduces a source of error in researches which use exposures of less than 10 m. sec.) Fig. 2 shows the 'build-up plus full-on' time for one of the Gerbrands fluorescent tubes when Sw_2 was set for 0.07 mfd. and the resistance at Sw_3 was varied within the range 0–120,000 Ω . The relationship is unmistakably linear. Thus, time-intervals other than those shown can easily be obtained by interpolation.

A LEVER-PRESSING DEVICE FOR FLUID REWARDS

By MERLE RIDGLEY and LAWRENCE I. O'KELLY, University of Illinois,
and JOHN L. FALK, Orange Park, Florida

While there are many satisfactory lever-boxes for the delivery of solid food pellets to rats and other small mammals, it has been more difficult to secure similar apparatus for measuring instrumental responses with fluid

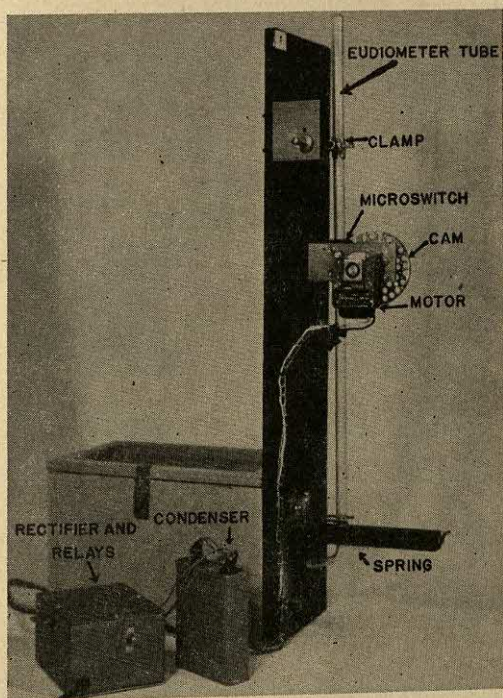


FIG. 1. PHOTOGRAPH OF DRINKING BOX

rewards. One of the major problems has been accurate measurement of the amount of fluid ingested. The apparatus described below has given reliable service and is capable of great flexibility in arrangement of specific operant and delivery components.

As shown in Fig. 1, the essential components are an eudiometer tube held on an upright support by a burette clamp whose shaft is free to

rotate, a cam driven by a synchronous motor, and a spring whose tension holds the eudiometer and its attached drinking tube against the cam edge. The animal activates the mechanism by pressing a lever; this closes a latching relay and starts the motor. Cam-rotation pushes the drinking tube into the box for a period of time determined by (a) the speed of the motor, and (b) the conformation of the cam. The latching relay insures that lever depressions during the drinking cycle are not registered on the attached recording device (although another circuit to a counter may be used to record *all* responses). As the cam rotates it depresses a microswitch which stops the motor at the end of a cycle.

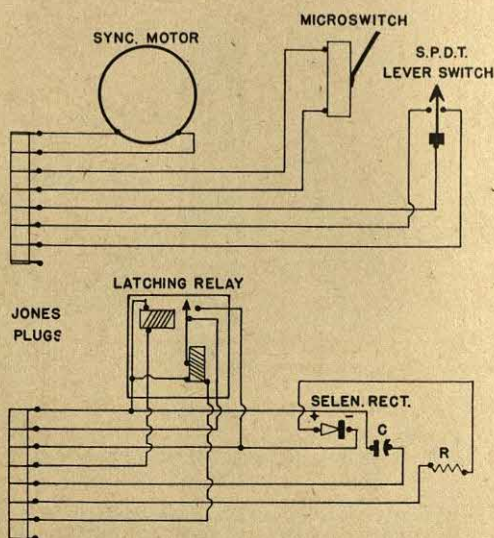


FIG. 2. WIRING DIAGRAM OF DELIVERY APPARATUS

Circuit values: selenium rectifier, 100 ma.; condenser, 10 mfd., 600-v. D.C., resistor, 200 Ω , 10 w.

The wiring diagram is shown in Fig. 2. The latching relay is activated by a condenser discharge. This makes it necessary for the animal to release the bar before a second response will activate the motor, thus preventing continuous cycling in response to passive pressure on the lever. If a record of the total number of responses is desired, a relay wired to Prongs 1 and 7 of the Jones Plug, and a counter or other recording device wired through the contacts of this relay may be used.

The sound of the relays and switching mechanisms are successfully damped by enclosing them in a blanket of glass-wool in a small box. In

practice we have operated six of the lever-boxes simultaneously. Motor and cam sounds may be masked by a fan or, as we have done in recent experiments, by a tape playback of actual drinking apparatus noise at a moderately high level of intensity.

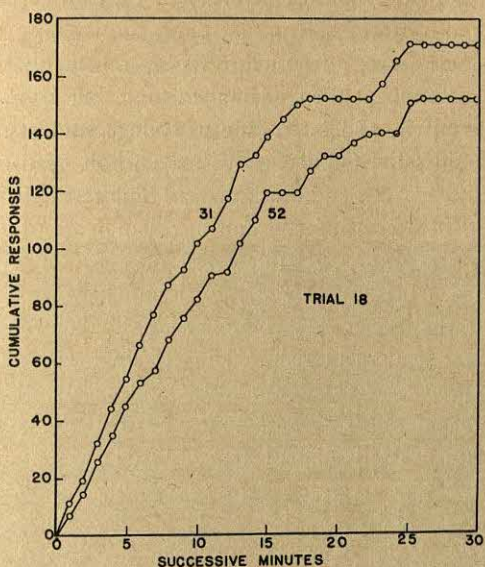


FIG. 3. DRINKING RECORDS ON THE 18TH TRIAL

On this trial Rat 31 drank 17.6 cc., 6.17% of its body-weight, and 0.103 cc. per response; Rat 52 drank 18.6 cc., also 6.17% of its body weight, at a rate of 0.122 cc. per response

Animals readily learn to operate the mechanism. The records, as the sample in Fig. 3 shows, are very similar in general form to those obtained with hungry animals and food reward in the conventional lever-boxes.

AN INEXPENSIVE VARIABLE-PATTERN MAZE

By WILLIAM F. DUKES, University of California at Davis

Many psychologists have attempted to design a maze that would be not only an efficient instrument for studying learning, but also an inexpensive one, easy to administer, and interesting to college students. These inventions include, besides the familiar stylus mazes, bolt head mazes,¹ single-path mazes,² finger mazes,³ and block mazes.⁴ The researcher or laboratory director, desirous of decreasing cost and increasing ease of administration, will, after careful scrutiny, probably reject most of the mazes described in the psychological literature and try to design his own. The invariable character of the traditional stylus maze makes it a one-situational problem; the electrical equipment required in certain mazes may make them too costly for large-scale usage; blindfolds and excessively tiring actions demanded by other mazes may be a source of irritation to the Ss.

Using the Tolman-Hall-Bretnall punchboard maze as a model,⁵ the present writer devised a maze that seems to meet the specifications outlined in the first sentence above. A piece of $\frac{1}{8}$ -in. masonite pegboard, 18 by 20 in., with holes $\frac{1}{4}$ -in. diameter and $\frac{3}{4}$ -in. apart, 46 golf tees (with points removed) to serve as pegs, 23 small squares of bright yellow paper, and an equal number of dark red paper were assembled. The pegs were inserted in pairs (in adjacent holes) in such a fashion that the lines described by them were in some cases horizontal, in some vertical, in others diagonal. In the pattern shown in Fig. 1, 23 such pairs were used. Beneath (on the underside of the masonite board) one member in each pair of holes, a yellow patch was pasted and beneath the other number a red

¹ Herbert Gurnee, A portable bolthead maze, this JOURNAL, 51, 1938, 405-406; C. W. Mann and W. O. Jewell, Jr., Configural aspects of human learning on the electrical maze, *ibid.*, 54, 1941, 536-545.

² W. A. Bousfield, The construction of a punch maze, *J. exp. Psychol.*, 35, 1945, 330-333. A. L. Rautman, A single-path punch maze for elementary experiments in learning, this JOURNAL, 61, 1948, 268-271.

³ R. G. Whisler, A universal flexible elevated finger maze, *J. gen. Psychol.*, 10, 1934, 482-484; J. Buel, A simple inexpensive finger-maze for experimental and classroom use, *J. exp. Psychol.*, 17, 1937, 192.

⁴ Shirley Morrison and T. W. Cook, An adjustable maze for visual and non-visual learning, this JOURNAL, 46, 1934, 141.

⁵ E. C. Tolman, C. S. Hall, and E. B. Bretnall, A disproof of the law of effect and a substitution of the laws of emphasis, motivation and disruption, *J. exp. Psychol.*, 15, 1932, 601-614.

one. The horizontal-diagonal-vertical character of the paired holes and the position of the colored papers were randomized. To avoid confusion about the pair membership of any peg, at least one hole was left vacant between adjacent pairs of pegs.

S's task is to start in the upper right corner, select one peg from each pair, remove it, note the color of the patch underneath, replace it, and move without hesitation to the next pair ('next' being indicated by chalked arrows), repeating the procedure. One color (yellow, for example) is designated as the 'correct' choice, and *S* is instructed that his goal is to learn to find his way through this series of choices exposing only this color. Recording choices is a simple matter.

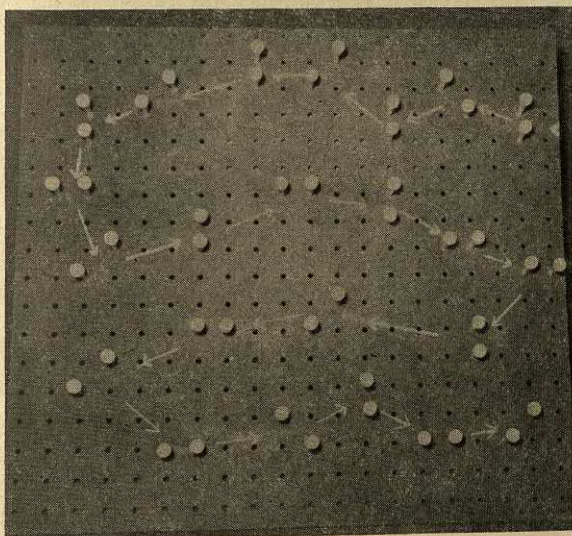


FIG. 1. AN INEXPENSIVE, VARIABLE-PATTERN MAZE

The maze as it appears to *S* is pictured in Fig. 1. This, of course, is but one of many learning situations possible with these materials. The experimenter may rearrange the pattern, lengthen it, shorten it, make a three-point choice problem or a linear maze, etc. While it is realized that the maze may no longer be considered an indispensable item of a psychological laboratory, this type of maze should be of value in laboratory situations which require that a large number of students work on the same exercise simultaneously and in research situations in which the same *Ss* are to be studied in a number of similar learning situations.

NOTES AND DISCUSSIONS

A SECOND TYPE OF SUPERSTITION IN THE PIGEON

When food is given to a hungry organism, any behavior in progress at the moment must be assumed to be reinforced by this event. When small amounts of food are repeatedly given, a 'superstitious ritual' may be set up. This is due not only to the fact that a reinforcing stimulus strengthens any behavior it may happen to follow, even though a contingency has not been explicitly arranged, but also to the fact that the change in behavior resulting from one accidental contingency makes similar accidents more probable. In an earlier experiment the automatic operation of a food-magazine every 15 sec. was found to induce hungry pigeons to engage in such ritualistic behavior as bowing, scraping, turning, and dancing.¹ In some cases the behavior was stable, in others the topography slowly changed; but in all cases superstitious effects survived indefinitely. Similar effects from the adventitious reinforcement arising from the presentation of discriminative stimuli have recently been observed.² Such effects must always be allowed for in designing experiments on complex behavior.

Accidental, but nevertheless effective, relationships may arise in the sensory control of operant behavior. For example, a stimulus present when a response is reinforced may acquire discriminative control over the response even though its presence at reinforcement is adventitious. Suppose, for example, that an organism is responding at a moderate rate on a variable-interval schedule of reinforcement, and let an incidental stimulus (*A*) occasionally appear for a brief period. Even though there is no explicit temporal relation between the appearance of *A* and the program of reinforcement, a response will occasionally be reinforced in the presence of *A*. For a brief period the frequency of such reinforcement may be appreciably greater than in the absence of *A*. An organism which is sensitive to slight differences in rate of reinforcement will form a discrimination; its rate of responding in the presence of *A* will become greater than in the absence of *A*. This might be called a positive sensory superstition. If, on the other hand, reinforcements happen to occur relatively infrequently in the presence of *A*, a discrimination will develop

¹ B. F. Skinner, 'Superstition' in the pigeon, *J. exp. Psychol.*, 38, 1948, 168-172.

² W. H. Morse, An analysis of responding in the presence of a stimulus correlated with periods of non-reinforcement, Unpublished doctoral dissertation, Harvard University, 1955.

in the opposite direction, as the result of which the rate in the presence of *A* will be relatively low—a sort of negative sensory superstition.

When an accidental contingency has produced a higher or lower rate of responding in the presence of an incidental stimulus, a second effect follows. If the rate has fallen in the presence of *A* (because reinforcements have been relatively infrequent), responses will be even less likely to be reinforced in the presence of *A*. In the limiting case no responses will be made in the presence of *A*, and no response, of course, reinforced. Moreover, reinforcements which are made available during *A* are not obtained because responses are not made. The first response following the withdrawal of *A* is then reinforced, and the discrimination is further strengthened. Similarly, when the rate is increased during *A* because of favorable accidental reinforcements, all reinforcements set up during *A* are likely to be obtained, and if the preceding condition commands a relatively low rate, some reinforcements set up at that time may actually be obtained after *A* has appeared, to strengthen the discrimination.

Both types of 'sensory superstition' have been demonstrated experimentally in the pigeon. The apparatus consisted of the usual experimental space $13 \times 22 \times 16$ in. A small translucent plastic plate was mounted behind a 1-in. circular opening at head height on one wall. It was lighted from behind by an orange 6-w. bulb. The pigeon pecked this disk to operate the controlling circuit. Food was presented for reinforcement in an opening below this key. Water was available in the space.

Three pigeons, two of which had previously been reinforced on other schedules, were placed on a variable-interval schedule of reinforcement with a mean interval of 30 min. The shortest interval between reinforcements was 1 min., the longest 59 min. Daily experimental sessions varied between 6 and 20 hr. in length. Body weight was maintained at approximately 80% of the *ad lib* weight, and a reinforcement consisted of access to mixed grain for 5 sec. The resulting performance was at a low mean rate of responding with some local irregularity. Against this base-line, an incidental stimulus consisting of a blue light projected on the key instead of orange was introduced for 4 min. once per hour. The schedule of occurrence of this stimulus was independent of the programming schedule. The rate of responding was recorded continuously in the usual cumulative curve. Brief downward movements of the pen marked reinforcements, and the pen remained down during the 4-min. period of the incidental stimulus, thus slightly displacing the record made in the presence of the stimulus without changing its slope.

Segments of records showing superstitious differences in rate under the

control of the incidental stimulus are shown in Fig. 1. Except for the portions *a*, *b*, *c*, and *d*, the curves are characteristic of the base-lines obtained on the schedule described above. Curves *A* and *B* are for a pigeon which showed sustained periods of negative superstition. The overall rate generated by the schedule is relatively high (of the order of 0.5 responses

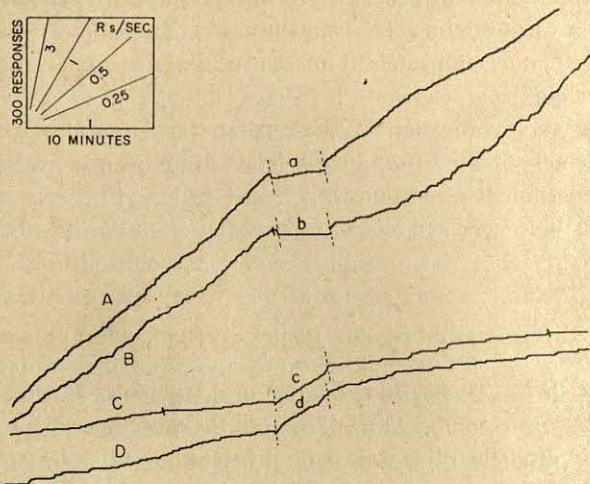


FIG. 1. SUPERSTITIOUS DISCRIMINATIVE CONTROL OF RESPONDING BY AN INCIDENTAL STIMULUS

The stimulus is presented for 4 min. once per hour, and is marked by the downward displacement of the recording pen (*a*, *b*, *c*, and *d*). Curves *A*, and *B* are segments of cumulative response-curves on a variable-interval schedule (mean interval = 30 min.) for a pigeon with a *lower* rate when the incidental stimulus is present than when it is absent—a 'negative superstition.' Curves *C* and *D* are segments for another pigeon on the same schedule with a *higher* rate in the presence of the incidental stimulus—a 'positive superstition.'

per sec.). Whenever the incidental stimulus appears, the rate drops to a low value or to zero (*a* and *b*). Records *C* and *D* are for another pigeon showing a positive superstition. The base rate is relatively low, and the rate in the presence of an incidental stimulus at *c* and *d* is clearly of a much higher order.

The direction of the superstition is not necessarily stable. In a long experimental session a positive superstition may give way to a negative form, or vice versa. Such changes are usually easily explained in terms of adventitious reinforcement or failure to receive reinforcement in the presence of the incidental stimulus. All three birds showed periods of both positive and negative superstition.

There are several arbitrary features of such an experiment. In the present case the incidental stimulus was present 1/15 of the time during

a session. A relatively shorter period would be less likely to receive reinforcements on a given schedule, and might be expected to produce negative superstition more frequently. At the other extreme, an incidental stimulus which occupied half the experimental session would presumably share so nearly equally in the reinforcements that there would be no substantial separation of rates. The schedule and the performance generated are also relevant in determining the frequency of adventitious reinforcement. Finally, the nature and intensity of the incidental stimulus also may have their effect.

Pending an investigation of these parameters, it may at least be said that incidental stimuli adventitiously related to reinforcement may acquire marked discriminative functions.³

Harvard University

W. H. MORSE

B. F. SKINNER

SOME EFFECTS OF INTERMITTENT SILENCE

Imagine that a monkey hits the keys of a typewriter at random, subject only to these constraints: (1) he must hit the space bar with a probability of $p(*)$ and all the other keys with a probability of $p(L) = 1 - p(*)$, and (2) he must never hit the space bar twice in a row. I wish to examine the monkey's output, not because it is interesting, but because it will have some of the statistical properties considered interesting when humans, rather than monkeys, hit the keys.

In the monkey's output we will find runs of i letters in a row, where $i = 1, 2, 3, \dots$, separated by single spaces. We will expect to find runs of length i with a probability

$$P_i = p(*)p(L)^{i-1}, \quad i = 1, 2, \dots, \quad [1]$$

so that the probability of a word of length i will decrease exponentially as i increases.

Now suppose that there are A different keys on the typewriter, excluding the space bar. Then the number of different possible words of length i must be A^i . Since the probability of any particular word must be P_i divided by the A^i different words of the same length, it must be equal to $P_i/A^i = p(*)p(L)^{i-1}A^{-i}$, on the assumption that all sequences of letters are equally probable. If we represent this quantity, the probability of a word of length i , by the symbol, $p(w, i)$, we can write it as:

$$p(w, i) = [p(*)/p(L)]e^{-i[\log A - \log p(L)]}, \quad [2]$$

³ This work was done under a grant from the National Science Foundation.

where w is a word of length i and everything else other than i on the right side is a constant.

Since there are A different keys available, there must be A words one letter long, $A + A^2$ words of length no greater than two, $A + A^2 + A^3$ of length no greater than three, etc. Therefore, we can write the general expression that there must be

$$\sum_{i=1}^k A^i = A(1 - A^k)/(1 - A), \quad (A \neq 1) \quad [3]$$

words of length less than $k + 1$.

Now suppose that we rank order all the different words with respect to length. The one-letter words will all get ranks between 1 and A , the two-letter words will all get ranks between $A + 1$ and $A(1 - A^2)/(1 - A)$, the three-letter words will all get ranks between $[A(1 - A^2)/(1 - A)] + 1$ and $A(1 - A^3)/(1 - A)$, etc. Thus the particular word, w , of length i , will be assigned an average rank, $r(w, i)$, given by the expression

$$\begin{aligned} r(w, i) &= [[A(1 - A^{i-1})/(1 - A)] + 1 + [A(1 - A^i)/(1 - A)]]/2 \\ &= A^i[(A + 1)/2(A - 1)] - [(A + 1)/2(A - 1)]. \end{aligned} \quad [4]$$

It will prove convenient to rewrite Equation [4] as follows:

$$[2(A - 1)/(A + 1)] \cdot [r(w, i) + (A + 1)/2(A - 1)] = e^{i \log A}. \quad [5]$$

Now we can combine Equations [2] and [5]. This is seen most easily if we first rewrite Equation [2] as follows:

$$p(w, i) = \frac{p^{(*)}}{p(L)} [e^{i \log A}]^{-[1 - \{\log p(L)/\log A\}]}$$

so that it becomes obvious how to substitute the left side of Equation [5]:

$$p(w, i) = \frac{p^{(*)}}{p(L)} \left[\frac{2(A - 1)}{A + 1} \left[r(w, i) + \frac{A + 1}{2(A - 1)} \right] \right]^{-1 - \{\log p(L)/\log A\}} \quad [6]$$

And this mess is what we wanted to derive. Actually, Equation [6] looks quite harmless after we define some new constants, substitute them in Equation [6], and write it as

$$p(w) = b[r(w) + c]^{-d}, \quad [7]$$

for now we have a simple relation between the probability of a word and its rank order. (Since probability decreases with length, once the words have been ordered with respect to increasing length, they are also ordered with respect to decreasing probability.)

Now suppose that we take the values $A = 26$ and $p(*) = 0.18$ from the English language and compute the constants:

$$b = \frac{.18}{.82} \left(\frac{50}{27} \right)^{-1.06} = 0.11, \quad c = \frac{27}{50} = 0.54,$$

$$d = 1 - \frac{\log 0.82}{\log 26} = 1.06$$

Then we have

$$p(w) = 0.11[r(w) + 0.54]^{-1.06} \quad [8]$$

If we wanted to round things off a bit, we might overlook the values 0.54 and 0.06 and so obtain the simpler expression,

$$p(w) = 0.11r(w)^{-1} \quad [9]$$

Equation [9] is very similar to a relation that is often called 'Zipf's Law.' Zipf found approximately $p(w)r(w) = 0.10$ in most natural languages.¹

Research workers in statistical linguistics have sometimes expressed amazement that people can follow Zipf's Law so accurately without any deliberate effort to do so. We see, however, that it is not really very amazing, since monkeys typing at random manage to do it about as well as we do.

Equation [7] is due to Mandelbrot.² In his original derivation, Mandelbrot minimized the average cost per unit of information. Later, Simon advanced a different explanation of Zipf's rule, one based upon the model of a 'pure birth process.'³ Simon's equation does not fit the verbal data as well as does Mandelbrot's, but it would seem to have the advantage of being derived from averaging rather than maximizing assumptions. The present argument should make it clear, however, as it was clear to Mandelbrot, that maximization is not essential to the derivation of Equation [7]. The assumption of maximization can be replaced by the assumption of random spacing.

In other words, Mandelbrot did more than derive Zipf's rule. He went on to show that the random placement of spaces which leads to Zipf's rule is actually the optimal solution. Our monkeys are doing the best possible

¹ G. K. Zipf, *The Psychobiology of Language*, 1935, 44-47; *Human Behavior and the Principle of Least Effort*, 1949, 19-55.

² B. Mandelbrot, *Jeux de communication*, *Publ. Inst. Stat. Univ. Paris*, 2, 1953, 1-124; Simple games of strategy occurring in communication through natural languages, *Trans. I. R. E.*, PGIT-3, 1954, 124-137. Also abstracted briefly in G. A. Miller, *Communication*, *Annu. Rev. Psychol.*, 5, 1954, 401-420.

³ H. A. Simon, On a class of skew distribution functions, *Biometrika*, 42, 1955, 425-440.

job of encoding information word-by-word, subject to the constraints we imposed on them. If we were as smart as the monkeys, we, too, would generate all possible sequences of letters and so make better use of our alphabet. Instead, we use only a small fraction of the possible letter sequences. In our behalf, however, it should be added that the consequent redundancy serves as insurance against errors. Since it is possible to regard redundancy as reducing the effective size of the alphabet, it should be noted that variation in A has little effect in Equation [6] relative to the effects produced by changes in p^* . Thus if redundancy reduced us to the equivalent of three letters, Equation [6] would change to

$$p(w) = \frac{p^*}{p(L)} [r(w) + 1]^{-[1-2.2 \log p(L)]}, \quad [10]$$

whereas, at the other extreme where $A \rightarrow \infty$,

$$p(w) = \frac{p^*}{2p(L)} [r(w) + 0.5]^{-1}. \quad [11]$$

Mandelbrot has found that the parameter d varies around 1.2, which corresponds rather well to Equation [10], when $p(L) = 0.82$, for then $d = 1 - 2.2 \log 0.82 = 1.187$. Similarly, an alphabet of $A = 3$ letters plus a space corresponds fairly well with Shannon's estimate of the redundancy of printed English.

It seems, therefore, that Zipf's rule can be derived from simple assumptions that do not strain one's credulity (unless the random placement of spaces seems incredible), without appeal to least effort, least cost, maximal information, or any branch of the calculus of variations. The rule is a simple consequence of those intermittent silences which we imagine to exist between successive words.⁴

Harvard University

GEORGE A. MILLER

THE USE OF FOREIGN LANGUAGES BY PSYCHOLOGISTS, CHEMISTS, AND PHYSICISTS

A previous note demonstrated that psychologists do not draw from other national literatures in proportion to the distribution of materials by language and country.¹ It is interesting to test this same hypothesis concerning two sister sciences—chemistry and physics. The present data for

⁴ This note was prepared under Contract AF33(038)-14343 and appears as report number AFCRC TN 56-60, ASTIA, Document No. AD 98822.

¹ C. M. Louttit, The use of foreign languages by psychologists, this JOURNAL, 68, 1955, 484-486.

psychology were taken from the previous article covering the year 1952. For chemistry and physics, the first 1000 consecutive literature citations were taken from the 1952 volumes of the following journals: *Journal of the American Chemical Society*; *Chemische Berichte*; *Annales de Chemie*; *Physical Review*; *Zeitschrift für Physick*, and *Annales de Physique*.

TABLE I
LANGUAGE OF CITATIONS IN PSYCHOLOGICAL, CHEMICAL, AND PHYSICS JOURNALS
Percentage of citations in

Language of publication		Percentage of citations in				
		English	German	French	Other	N
Psychology	English	92.5	5.2	0.7	1.6	879
	German	2.7	91.1	6.2	0.0	112
	French	25.6	7.4	64.0	3.0	203
	Total	72.7	13.6	12.0	1.7	1194
	Psychological Abstracts	77.0	5.5	8.0	9.0	
Chemistry	English	79.4	11.9	4.7	4.0	1000
	German	22.8	64.0	6.2	7.0	1000
	French	36.7	27.9	29.1	6.3	1000
	Total	46.3	34.6	13.3	5.8	3000
	Chemical Abstracts	54.7	11.1	9.2	25.0	
Physics	English	86.2	5.9	2.0	5.9	1000
	German	35.3	58.1	1.4	5.2	1000
	French	49.8	15.3	30.7	4.2	1000
	Total	57.1	26.4	11.3	5.1	3000

While it is impossible exactly to define the universe from which citations might be drawn, for psychology and chemistry there are two possibilities. The first is the total number of references cited (see Table I). A second for psychology consists of the entries in the 1952 volume of *Psychological Abstracts*, also given in Table I. The second possibility for chemistry is an analysis of the *Chemical Abstracts* for 1949 performed by Boig and Howerton and presented in a series of two articles.² The percentages reported in Table I were computed from data in Table 3 in each of these publications. Incidentally, the very large percentage of 'Other' languages is largely made up of Russian, 14.6% of the total. For physics only the total number of references can serve as an estimate of the universe.

In all three sciences, it is clear that the hypothesis of proportional citation can be rejected regardless of how the universe is defined. In all three sciences, the English and German scientists cite their own language far in excess of its relative proportion in the universe. While French

² F. S. Boig and P. W. Howerton, History and development of chemical periodicals in the field of organic chemistry: 1877-1949, *Science*, 115, 1952, 25-31; History and development of chemical periodicals in the field of analytical chemistry: 1877-1950, *ibid.*, 115, 1952, 555-560.

scientists do likewise, they rely less upon their own language and cite English and, especially in chemistry, German more heavily.

The differences between the three sciences are no less striking. English and German psychologists cite their own languages almost exclusively. English and German chemists and physicists cite their own languages less frequently than psychologists, tending to cite each other's language more often. French psychologists use English citations, but not nearly as frequently as French chemists and physicists do. The French chemists and physicists are most notable in that they are the only groups that cite languages other than their own a majority of the time (approximately 70% in each case). All of the above statements that imply a significance test have as a basis a Chi-square test significant at the 1% level of confidence.³

Wayne State University

C. M. LOUITT

WORD-LENGTH AS A FACTOR IN DIFFERENTIAL RECOGNITION

The term 'differential word-recognition' has been used to describe the finding that tachistoscopic recognition is greater for words placed to the right of a fixation-point than for words placed at the same distance to the left of it.¹ Mishkin and Forgays explain this result as a consequence of special training of the left hemi-retinas for word-recognition in the course of experience in reading English.² On this basis Orbach and others have attacked the doctrine of equipotentiality.³

If the doctrine of equipotentiality is to be rejected, the differential recognition of words should not be capable of being explained in terms of differences in acuity or differences in importance of the parts of words. Huey has argued that the first letters of words are very important in recognition.⁴ When words are to the right of fixation, their first letters fall closest to central vision (in an area of greater acuity than the last letters); the reverse is true of words on the left. If Huey's contention is correct,

³ Dr. Louttit had collected and partially analyzed these data shortly before his untimely death. Robert L. McCornack completed the analysis and wrote this note, and must be held responsible for any errors that may have been introduced due to his unfamiliarity with all the details of the study.

¹ D. G. Forgays, The development of differential word-recognition, *J. exp. Psychol.*, 45, 1953, 165-168.

² Mortimer Mishkin and D. G. Forgays, Word-recognition as a function of retinal locus, *ibid.*, 43, 1952, 43-48.

³ Jacob Orbach, Retinal locus as a factor in the recognition of visually perceived words, this JOURNAL, 65, 1952, 555-562.

⁴ E. B. Huey, *The Psychology and Pedagogy of Reading*, 1908, 96-99.

words to the right of fixation should be better recognized than words to the left. Furthermore, longer words should show greater differential recognition since their first letters would be relatively closer to central vision for positions to the right of fixation. Such a result, however, would not be predicted by Mishkin and Forgays' experiments.

The purpose of the present note is to report the results of experiments testing this explanation; namely, the relative importance of the first letters of words. The null hypothesis was that seven-letter words would not show significantly greater differential recognition than three-letter words.

Previously established criteria of age, visual ability, education, and eye-dominance were used to select 81 Ss (60 men and 21 women) from among 120 university students who volunteered for the experiment. A fixed head-rest placed S's eyes 2 m. from a white screen. A Keystone Overhead Tachistoscope was so placed that the projected image of tall letters in pica type averaged 2 cm. in height. Words projected on the screen were centered either 10.6 cm. to the right or 10.6 cm. to the left of the fixation-point. The words used were selected from the Thorndike-Lorge lists; there were 10 three-letter words and 10 seven-letter words.⁵

S was instructed to look at the fixation-point and to report each word exposed. E recorded each response and watched S's eyes for gross eye-movements. Each S viewed 5 seven-letter words and 5 three-letter words

TABLE I

MEAN NUMBER OF WORDS RECOGNIZED FOR EACH WORD-LENGTH AND WORD-POSITION

Word-length (in letters)	Word- position	Mean No. of words recognized	SD	t
Three	right	2.53	1.56	0.76
	left	2.37	1.22	
Seven	right	3.51	1.42	7.56*
	left	1.89	1.42	

* Significant at the 1/10% level of confidence.

to the right of the fixation-point and the same number on the left; the order of presentation was counterbalanced. A given word appeared on the right for one S and on the left for the next. Exposure-times ranged from 10 to 100 m.sec., and monocular and binocular vision were counterbalanced.⁶

⁵ E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, 1944.

⁶ In the original study several such factors were investigated. Since these two were not found significantly related to the differential recognition of words, results were appropriately combined.

Analysis of variance indicated a significant (1/10% level of confidence) interaction between word-length (three or seven letters) and word-position (right or left of the fixation-point). The accompanying table shows the direction of this interaction; significant differential recognition was found for seven-letter words but not for three-letter words. The null hypothesis was rejected at the 1/10% level of confidence.

The prediction from Huey's argument was supported. This result provides evidence for his explanation of the differential recognition of words set forth above. Since the doctrine of equipotentiality would allow such an explanation, the doctrine need not be rejected on the basis of previous studies in this area.

University of Florida

JOSEPH R. MELVILLE

AN INTEROCULAR COLOR-EFFECT

Several years ago I accidentally discovered a phenomenon which seems to have important implications for the theory of color-vision. Though the conditions under which it may be observed are very simple, I have never found it described or mentioned in the literature.

The following description is in terms of operations which the reader may perform with materials close at hand. A paper of some highly saturated color (*e.g.* red) is held before one eye (*e.g.* the right) at an angle such that it is well-illuminated and fills most of the field of that eye. The opposite eye remains open and moves normally. After some 20-60 sec., the paper is removed and the observer looks at a white or neutral gray surface, closing first one eye and then the other in fairly rapid alternation. With the right or stimulated eye, the surface is seen as greenish; this is merely to say that a negative after-image appears. With the left or unstimulated eye, however, the surface is seen as slightly *pink*; *i.e.* the inducing hue appears in reduced saturation.

Lights with appropriate filters may of course be substituted for colored papers. Using papers, I have found red and yellow to give more striking results than other hues. Presumably the inducing hue appears to the unstimulated eye as the result of a double-reversal process, the first reversal being that of the negative after-image, the second being an effect of contrast with the after-image. The after-image-reversal clearly involves an opposition of complementaries in the system of a single eye, and may well occur in the retina. The contrast-effect, on the other hand, is *interocular*, and therefore implies an opposition of complementaries in some region

central with respect to the optic chiasma. Possibly it occurs in a mechanism the normal function of which is to maintain constancy of hue.

The effect merits a more systematic investigation than I have been able to give it.

Center for Advanced Study
Stanford, California

FRED ATTNEAVE

A GROUP-METHOD FOR THE STUDY OF FIGURAL AFTER-EFFECTS

It is the purpose of the present note to report the availability of a group method for the study of figural after-effects. The technique facilitates the collection of experimental data, and may prove valuable as a simple demonstrational experiment for introductory courses.

The figures employed were adapted from Köhler and Wallach.¹ They were drawn in pencil on white bond paper. The 'wings' of the Muller-Lyer *I*-figure were $1\frac{1}{2}$ in. long and $2\frac{1}{2}$ in. apart at the apices, with the line often seen joining the wings omitted. The apparently shorter side of the illusion was on the right. The *T*-figures were always pairs of squares: the left-hand square was always 1 in. on a side, while the right-hand square varied in size from $\frac{7}{8}$ to $1\frac{1}{8}$ in., in $\frac{1}{16}$ -in.-steps. The centers of the squares of each *T*-figure were $2\frac{1}{2}$ in. apart. The figures, which also contained appropriate fixation-points, were magnified $4\frac{1}{2}$ times on a portable motion-picture screen by a Squibb-Taylor Model IS-3 opaque projector.

The *Ss* were instructed that they would be shown some simple geometric figures (the *T*-figures) and that they were to judge, as quickly as possible while looking at the fixation-point, whether the square on the right was larger, smaller, or equal in size to the square on the left. A control group made repeated judgments of the *T*-figures, while for an experimental group, appropriately instructed, the *I*-figure preceded each *T*-judgment—for 2 min. before the first and for 20 sec. before each of the subsequent judgments. The order of presentation of the *T*-figures was the same for both groups and was randomized with the restrictions (1) that the first judgment of the experiment was of the *T*-figure having squares of equal size, and (2) that each of the *T*-figures was shown once during each block of five trials; the latter restriction was imposed in an attempt to insure

¹ Wolfgang Köhler and Hans Wallach, Figural after-effects, *Proc. Amer. Phil. Soc.*, 88, 1944, 278.

enough comparability of halves of the test that a coefficient of reliability could be computed. Forty judgments of the five *T*-figures (eight per figure) were required of all *Ss*. The split-half reliability for 41 *Ss* was 0.78. The figural after-effect was evidenced in terms of significant differences between the judgments of the two groups.

Indiana University, Jeffersonville Center

JOHN C. JAHNKE

FIFTY-THIRD ANNUAL MEETING OF THE SOCIETY OF EXPERIMENTAL PSYCHOLOGISTS

The fifty-third annual meeting of the Society of Experimental Psychologists was held in Memorial Hall at Harvard University April 4-5, 1957. Edwin B. Newman served as chairman of the meeting. The following members attended: Bartley, Beebe-Center, Boring, Bray, Chapanis, Dallenbach, Estes, Gagne, Garner, Gibson, Graham, Grant, Harlow, Hovland, Humphreys, W. A. Hunt, Kennedy, Licklider, Lindsley, Marquis, Melton, G. A. Miller, N. E. Miller, Morgan, Mueller, Nafe, Neff, Newman, Pfaffman, Pratt, Riggs, Schlosberg, Skinner, K. U. Smith, Solomon, Stevens, Underwood, Volkmann, and Wickens.

J. P. Egan, D. H. Lawrence, and F. Ratliff were elected to membership. The Society now has 81 members.

Scientific sessions were held the afternoon and evening of April 4 and the morning and afternoon of April 5. Members presented informal reports of research in progress.

At the annual dinner on April 4, the Warren Medal and Warren Medal Award for 1957 were presented to Lorin A. Riggs for "his extraordinarily fruitful techniques of recording electrical potentials from the human retina and for his many significant contributions to systematic knowledge and theory of vision that these techniques have made possible."

The dinner was also made the occasion for the presentation to Harvard University of the library fund to honor Edwin G. Boring who becomes Edgar Pierce Professor of Psychology Emeritus in June 1957. The fund was made possible by contributions of students and associates of Professor Boring. President Pusey accepted the fund for Harvard.

The Society received an invitation to hold its 1958 meeting at Northwestern University and its 1959 meeting at Johns Hopkins. Both invitations were accepted. Thursday and Friday of the first week in April were chosen as tentative dates.

University of Chicago

W. D. NEFF

THIRD ANNUAL MEETING OF THE SOUTHEASTERN PSYCHOLOGICAL ASSOCIATION

The Southeastern Psychological Association held its third annual meeting March 28-30, at Nashville, Tennessee, with Fisk University, George Peabody College, and Vanderbilt University as sponsors. The program was made up of an invited address "Psychology and mental health" by Fillmore H. Sanford, Joint Commission on Mental Illness and Health; 3 presentations of programmatic research, a workshop, 5 symposia, and 34 papers. The presidential address entitled "Motivation to high achievement" was presented by Nicholas Hobbs. Officers for 1957-58 are: President, Edward E. Cureton; President-Elect, W. N. Kellogg; Secretary-Treasurer, M. C. Langhorne; Members-at-Large on the Executive Committee, Dorothy C. Adkins, C. H. Calhoun, and Eliot H. Rodnick. The 1958 meeting will be held at Atlanta, Georgia, April 27-29.

Emory University

M. C. LANGHORNE

BOOK REVIEWS

Edited by GEORGE L. KREEZER, Washington University

On Human Communication: A Review, a Survey, and a Criticism. By COLIN CHERRY. Cambridge, The Technology Press of Massachusetts Institute of Technology, and New York, John Wiley & Sons, 1957. Pp. xiv, 333.

There are so many different ways to look at 'human communication' that these words in the title of a book only hint at what will be found between its covers. An anthropologist would develop the title very differently from, say, a philosopher or a librarian. It is important to know, therefore, that the author of his volume is an electrical engineer. More precisely, he is Reader in Telecommunication, Imperial College, University of London, which means that his inspiration springs primarily from the recent development of a mathematical theory of communication in the field of communication-engineering.

To label this as a book in electrical engineering, however, would be a sad mistake. One of Colin Cherry's most valuable contributions to this new theoretical outburst has been his gift of perspective, historical, philosophical, and scientific. He has the British talent for lucid exposition, and his lectures have captured the imagination of audiences on both sides of the Atlantic. Now he has organized and rewritten those lectures into a book which will serve as an admirable introduction to a series of more technical works by various authors ("Studies in Communication") that is planned for the years ahead. With this introductory role in mind, Cherry has made a sincere effort to relate the new, engineering version of communication to the more traditional problems of anthropology, linguistics, philosophy, and psychology. Although the result is less spectacular than Norbert Wiener's *Cybernetics*, the reader of Cherry's book will have a clearer picture of the field as it exists today. In spite of Wiener's virtuosity, his enthusiasm for the assets of the new theory often made it difficult to separate deposits from promissory notes. Now, nine years later, it is possible for Cherry to adopt a more disinterested attitude and to distinguish the many substantial accomplishments from the vast array of questions that remain to be answered.

Those who heard or read Cherry's lecture on the history of the theory of information will be pleased to find this material, which ranges from ancient Egyptian inscriptions to pulse-code modulation, incorporated in the second chapter. The analysis of speech-signals, which is close to one of Cherry's own research-interests, also has a chapter of its own. Chapter 5, "On the Statistical Theory of Communication," is an excellent introduction to the ideas of Hartley, Shannon, Wiener, Fano and other information-theorists, and Chapter 6 represents a valiant effort to relate the various approaches to information-measurement in terms of the familiar trichotomy of syntactics, semantics, and pragmatics.

Of most direct interest to psychologists, however, is the final chapter "On Cognition and Recognition." It is perhaps unfair to complain that this chapter does not

come off, for the aim was to review, survey and criticize, not to create a new psychological theory. Yet I cannot avoid the feeling that this essay is more a collection of odds and ends that had come to the author's attention at various times than a careful appraisal of cognition and recognition in the light of communication-theory. It is, in some sense, a compliment to Cherry to have hoped that he might have accomplished more with so difficult a psychological topic. "Leibnitz," to quote the opening sentence, "was the last man to know everything."

This is a book you can safely recommend as a good introduction to the mathematical theory of communication. The publishers and the public will indeed be fortunate if the rest of the series meets the high standards Cherry has set.

Harvard University

GEORGE A. MILLER

Annual Review of Psychology, Volume 8, 1957. Edited by Paul R. Farnsworth; Quinn McNemar, Associate Editor. Palo Alto, Annual Reviews, Inc., 1957. Pp. ix, 502.

In the preface of Volume 1 of the *Annual Review of Psychology* there is a clear description of the intended audience. This audience was to be primarily graduate students and teachers, and persons outside the field of Psychology. In general, Volume 8 addresses itself to this audience very well. The various reviewers are not equally successful in their efforts, nor do they perceive their tasks in exactly the same way. Some strive for relative completeness, others apply rather arbitrary criteria for selecting material for review. Some of these differences will be evident in the brief survey of the individual chapters to which we now turn, leaving our general comments for the end.

There is a group of chapters which refer to topics which may, for want of a better term, be included in "physiological psychology." Three of these chapters are concerned with sensation, and they differ radically in scope and conception. Vision (Pickford) aims at, and achieves, complete coverage of the psychological literature. Physiological and biochemical material is omitted, but there is a refreshing coverage of non-English sources. The chapter is so completely non-evaluative and non-selective as to make rather dull reading. The chapter on hearing (Lawrence) is almost exactly opposite. Here the emphasis is on anatomical, physiological, and biophysical topics. Perception, including communication, is almost ignored. But the chapter is well-organized and very readable. Somesthesia (Oldfield) does not come off as well. A large part of the psychological work is ignored, pain is poorly represented, and the temperature senses, kinesthesia, and the vestibular senses are scarcely mentioned. Since the previous review of somesthesia was largely a polemic (Weddell, Volume 6), and since Oldfield devotes much of his space to the same argument, somesthesia has not been given an adequate general review since Ruch's chapter in 1953.

Of the other "physiological" chapters, the one entitled "Physiological Psychology" (Stellar) is an enthusiastic, very selective discussion of articles considered significant, largely from neurophysiology. Biochemistry, pharmacology (except for tranquilizers), and endocrinology are not discussed. Only when writing of the tranquilizers does Stellar become truly critical. Comparative Psychology (Bindra) does stay close to the comparative approach, and is well-organized and interesting. Given the task set for himself by the reviewer, coverage is good. But the omission of all comparative sensory and "higher mental process" material limits the contribution. In the

chapter on Special Disabilities (Meyerson), we find an intelligent, critical discussion of blindness, deafness, and crippling. Possibly more material could have been covered had the author not chosen to devote so much space to (usually justified) general complaints. The chapter on Abnormalities of Behavior (Wright) is medically oriented, and the organization has little psychological significance. There is a generally uncritical approach, exemplified by the acceptance of standard diagnostic categories without comment. Some important material is omitted (for example, organic brain states are represented largely by chemical treatments for feeble-mindedness), and the exclusion of foreign references is most unfortunate in this field.

A second group of chapters may be described loosely as relating to "personality." The chapter on Personality (Eriksen) is well-integrated and evaluative. The criterion for inclusion of material is not, however, obvious. The physiological psychology of motivation and the biochemistry of stress seem more related to the reviewer's interests than to the topic at hand. Individual Differences (Jenkins and Lykken) is the toughest-minded chapter in the book, and comes very close to being an intellectual *tour de force*. The emphasis is upon methods, principles, and research techniques. Positive suggestions abound. The chapter has excellent organization, much insightful criticism, and is very well written. The chapter on Assessment (Wittenborn) is largely non-critical, and gives the impression of lack of integration. The area is not clearly defined. The emphasis is upon projective techniques, and most psychometrics is omitted. Psychotherapy (Winder) is given a surprisingly hard-headed, but optimistic treatment. One gets the impression of a well-organized coverage of a poorly-organized field. The chapter on Counseling (Shaw) identifies counseling with psychotherapy so completely as to define much of counseling out of existence. The hypotheses and concepts rampant in psychotherapy are not given critical treatment.

Two chapters may be said to relate to broader problems and techniques. The chapter on Learning (Walker) gives a critical and broad treatment of this important topic. The chapter is well-organized, covers many problems not always included, and represents a broadening conception of the field. Walker is intelligently critical of the popular band-wagons. The treatment of Statistics (Gardner) is judicious and neatly organized. Emphasis is upon assumptions and the refinements of older techniques. There is nothing on psychometrics (nor is it adequately treated elsewhere), and no reference to the developing importance of high-speed computers as determiners of research design.

The last chapters to be considered comprise the "special fields." Industrial Psychology (Katzell) is a very good overview, with excellent organization and evaluation. Actually, coverage leans toward "personnel" rather than "industrial" psychology so that such topics as consumer testing, quality control, and efficiency get rather short shrift. Nevertheless, the chapter is a valuable contribution to the topics covered. Social Psychology (Cartwright) is given broad treatment. The chapter provides a good idea of what is going on in the field, although the organization and criticism seem somewhat spotty. Educational Psychology (Keller and Corcoran) is given a well-organized non-critical overview. The conclusions appear to be sound, and emphasize the current preoccupation with applied problems. The chapter on Developmental Psychology (Inhelder) is difficult to read because of stilted language. Apparently, most of the current American work does not fit into the reviewer's theo-

retical position, which is given free play, but it is interesting to have the European coverage. There is some useful and insightful criticism, although a tighter organization might have been helpful.

So much for the details! What can be said in general? We may begin by reiterating our comment that Volume 8 meets the objectives of the review series rather well. Some additional comments are needed, however. First, the overlapping of material from one chapter to another turns out to be of minor proportion. Where overlapping occurs it seems adequately justified by the nature of the material. Second, it may be asked whether or not the space is adequate for proper reviews. We feel that it is, and lack of coverage of truly significant developments is a reflection on the skill of the reviewer. In this regard, it would be well for reviewers to omit lamentation about lack of space and how many articles had to be read, and use the same number of words for more constructive purposes. Third, the reviews are largely restricted to the English-language literature, and we must face our own provincialism. Perhaps the future holds the adoption of a single scientific language, and a consequent reduction in bibliographic burden. The staggering number of articles to be read and the statistical and other important tools which the present-day psychologist must have in his scientific armamentarium, combine to make real proficiency in other languages largely a thing of the past. Fourth, the announcement that hereafter titles of articles will be included in the references (as they are in a few chapters this year) is welcome. Even more welcome would be the adoption of the form of citation used in this JOURNAL and in those of the American Psychological Association. Fifth, reviewers should be chosen from year to year to represent diverse points of view. The choices in the field of hearing have been noteworthy in this respect. Also, persons with crippling biases should be avoided, since the stated purposes of the *Review* preclude polemics. And we would like to suggest that in many instances the bright young man does a better job than his academic senior who is likely to be both too dedicated to an opinion and too busy to do the job well.

We cannot close this review without asking whether or not we can get some impression of where psychology is going by reading these reviews. It seems to us that if Volume 8 reflects reasonably well the present state of psychology, four general trends may be detected. (1) More and more mathematical sophistication is required to comprehend current research. This is true not only for statistics, but for the mathematical models which are beginning to appear. (2) More and more physiological material is appearing, and the "empty organism" does not dominate the field. (3) Methodological sophistication is increasing, but a straight positivist approach is waning. (4) There are no big hypotheses, no new integrating principles, and the dimensions of description are still not satisfactory. Perhaps nothing corresponding to the periodic table or organic evolution is possible in our field. At any rate, it does not appear in Volume 8 of the *Annual Review of Psychology*.

University of California
Los Angeles

MARGARET H. JONES
F. NOWELL JONES

Vocational Interest Measurement: Theory and Practice. By JOHN G. DARLEY and THEDA HAGENAH. Minneapolis, University of Minnesota Press, 1955. Pp. xvii, 279.

Darley and Hagenah have written a thoughtful volume upon interest-measurement and its implications which should be of interest to psychologists at large even

though it is intended primarily for the counsellor. They deal with (1) procedures and results of standardizing and validating the Strong Vocational Interest Blank, and procedures for interpreting the scores, (2) empirical studies of the relationships among interest-scores and various other measures of personality, and (3) theories concerning the origin and development of interests, and of their relationship to variables of motivation, learning, and socialization. The restriction to the Strong Blank is not so limiting as might first appear, since there is strong justification for Darley and Hagenah's statement that "we know of no other psychometric device that has been subjected to such rigid scrutiny." (p. 56)

The first of the above topics is dealt with in three chapters. Of these, one is an excellent summary of the methods by which the Strong Blank was developed and validated. The other two form a sequel to Darley's earlier monograph (1941) in which he presented his method of analyzing Strong scores in terms of "primary," "secondary," and "tertiary" patterns. More recent normative studies for the method, based upon the later forms of the Strong Blank, are summarized here in some detail. There is also an extended chapter on the clinical interpretation of these patterns, illustrated by a series of case histories.

The treatment of this first topic is highly successful and leaves the reader with a sense of the practical value of interest-measurement in college counselling. The research available upon the other two topics is summarized as clearly and carefully as the first, but there are serious limitations in the material itself which make the research difficult to evaluate. The authors have chosen to present the results at their face value, however, without looking critically at the underlying methodological or conceptual problems. This orientation is understandable since their book seems to be directed primarily at the practical user of interest scales.

What follows is to be regarded, therefore, not as a criticism of Darley and Hagenah, but as a statement of certain deficiencies in the present state of the field which they are summarizing. These deficiencies need to be stated in order to characterize the status of theory in the field.

An impressive amount of research is summarized in the chapter dealing with the relation of interest-scores to other measures of personality. This reviewer was left, however, with an uncomfortable impression that the interpretations and conclusions move dangerously far from the raw data and from the concrete behavior, beliefs, and perceptions of the subjects.

Interest-scores are themselves derived in complex ways from answers to a questionnaire. The personality variables with which they are being correlated also are derived from answers to questionnaires. Overlooking the fact that there is overlapping in the content of the questionnaires themselves, there is still the possibility that the correlations may be a function of the way in which questionnaires are answered, rather than representing relationships which could be observed in interests and traits measured in some other way.

The interpretation of the interrelationships is further complicated by the fact that each study involves the working out of a large number of correlations on a single group of subjects, and that the answers to the same questions on the Strong Blank are used repeatedly in working out the various scores. The use of a common sample of subjects for studying a number of interrelationships is a feature of a large number of fields of research, and seems to be generally overlooked in applying significance

tests to the results. The use of a common set of questions for getting a number of different scores is a less widespread but also commonly ignored problem.

To take a single very crude example, it is not at all surprising that the "occupational level" scores correlate with various scores for specific occupations, if we look at the way in which scores for occupational level are developed. To describe these relationships as though they represented relationships between independently measured characteristics of the subjects is certainly misleading. On the other hand, to investigate these relationships is not redundant, since they cannot be predicted completely from the standardization procedure. It might turn out, for example, that the items which differentiate occupational levels are of relatively low weight in determining specific occupational scores. Thus such investigations have a definite use, but we suggest that they should be treated as studies on a quite different level from those in which different methods of measurement, based upon non-overlapping information, are used. In fact, there is a need here for a careful theoretical analysis of the methods of interpreting this kind of information.

The theoretical portion of the book is concerned with very broad questions, rather than with the kind of technical problem which we have mentioned above. Darley and Hagenah raise questions concerning the extent to which interest patterns are learned, how interests are related to ability, to self-concepts, social factors, and the like. The authors admit that there is a large gap between the empirical results which they have summarized and the theoretical discussions of these broad problems. Perhaps no one is now capable of closing the gap. At any rate, Darley and Hagenah have adopted the more modest aim of pointing out the problems "that others may succeed where we have failed."

This reviewer believes that more progress could be made if there were more attention paid to problems of definition. The question of whether interests are related to value-systems, needs, motivations, is meaningless unless these terms are explicitly defined. They can be defined in such a way that they are related by *definition* and no empirical studies are needed. A more fruitful approach would be to define different terms by means of different kinds of behavior, or behavior as observed in different ways. Then the problem of interrelation becomes an empirical one. Without definition, however, we cannot collect meaningful facts and interrelationships, or even define our problem. In Darley and Hagenah's discussions we are left to choose our own definitions of terms, and it is therefore difficult to decide exactly what issue is in question at any given point.

For the general psychologist, Darley and Hagenah's book makes clear that interest-measurement raises many issues about fundamental psychological problems—motivation, learning, personality development, and the like. It also shows that the interest-scale is a useful tool for research upon these problems as evidenced by the number of studies which the book summarizes. Some may also agree with this reader that the nature of this tool as a means of research must be more critically scrutinized.

Cornell University

T. A. RYAN

Personality in Young Children. By LOIS BARCLAY MURPHY and ASSOCIATES. New York, Basic Books, 1956. Vol. I, *Methods for the Study of Personality in Young Children*; Pp. xx, 424. Vol. II, *Colin—A Normal Child*; Pp. xii, 267.

Although an understanding of personality development during the preschool years

is of utmost practical and theoretical importance, there are surprisingly few good systematic studies of this subject. The deficiency probably stems in large measure from the paucity of adequate methods for appraising youngsters' complex needs and feelings, their perceptions of—and attitudes toward—themselves and the world about them, their frustrations and gratifications. Research workers in this field have customarily used teachers' or observers' ratings as basic data, although recently a few techniques developed in clinical work—doll play, interviews, and the Rorschach—have played a limited role.

These two volumes by Lois Murphy and her associates constitute a particularly welcome and significant contribution because they offer a diversity of novel approaches to the assessment of personality in young children. For over fifteen years, the authors made intensive case studies of the pupils attending the nursery school at Sarah Lawrence College. Throughout this period, they experimented creatively with many kinds of materials and situations which might provide new and deeper insights into young personalities. Their investigations have been frankly exploratory, many of them aimed less at finding answers to specific questions than at suggesting "a whole battery of questions [that the investigators] would not have known enough to ask."

Volume I describes the methods they found to be most effective in their work. These are grouped into four broad categories, each discussed by an expert: Lois Murphy describes *free methods* (miniature life toys, dough and cold cream, painting, and the Rorschach); L. J. Stone, techniques of *group play and readiness for destruction* (leadership games, destruction of balloons); Eugene Lerner, *active play techniques* (a series of "blocking" or frustration situations); and Evelyn Beyer, *observations in nursery school situations*. Volume II gives a thoroughly detailed and well-integrated report of the application of these techniques in an intensive study of *Colin—a normal child*. Simple toys (dolls, blocks, balloons), commonplace materials (dough, cold cream), and familiar games (school, house)—all intrinsically appealing to children—are ingeniously used as projective instruments. The entire battery provides opportunities to observe the child's reactions to structured and unstructured stimuli; to formal and informal situations; to adults and peers; and to frustrations and gratifications. The account of each technique is lucid and complete, including descriptions of the materials used, optimal conditions for testing, and instructions to the Ss. These, plus the many illustrative records and original protocols, make Volume I an excellent test-manual.

To most readers the most impressive aspect of the work will be found in the insightfulness of the interpretations, which reflect the investigators' deep understanding of and empathy with children, as well as their highly developed sensitivity to nuances of behavior and verbalization in the test-situations. Despite the difficulty of delineating precisely the bases for clinical inferences, these authors, with a few exceptions, are able to communicate their systems of analysis of data and much of the rationale underlying their interpretations. The numerous suggestions which stimulate sensitivity to subtleties of response and the discussion of the influences of familial, cultural, and situational factors on performance make profitable reading for all clinicians and students of child psychology. Many of the reports included, e.g. Beyer's observations in the nursery school, are models of psychological writing, combining detailed and precise observations, clearly reasoned interpretations, and fine prose style.

From this reviewer's point of view, there are two important deficiencies in the work. First, the problem of the validity of the interpretations is handled much too casually. There are a few anecdotal references to the relationship between test-results and mothers' or nursery school teachers' reports of behavior, but no systematic studies of validity are reported. Considering the wealth of available data on the Ss' personality and behavior during preschool and subsequently, it would probably not be difficult to carry out such studies. Secondly, the work lacks theoretical unity. All the investigators have been influenced by psychoanalytic theory, but they differ widely in the extent to which this affects their observations and interpretations. Only a few of them make their own conceptual schemes explicit, and there is no attempt to develop a unified conceptual framework. This shortcoming may seriously limit the usefulness of the test-battery, for without a consistent theoretical orientation it is difficult to integrate the data and interpretations derived from different techniques.

Despite these limitations the two volumes are thoughtful and stimulating. Clinicians can use the instruments qualitatively—as Murphy and her collaborators did—to achieve deeper insights into their child patients. In addition, the methods permit discriminating and thorough personality description and analysis which will facilitate the systematic study of personality development. A few child psychologists already have used them profitably in their research. These books are likely to convince others of their fruitfulness.

University of California

PAUL MUSSEN

Measurements of Mind and Matter. By G. W. SCOTT BLAIR. New York, Philosophical Library, 1956. Pp. 115.

This little book is an attempt to dispel some of the snobbishness among scientists concerning the primacy of so-called fundamental measurements in physics. Scott Blair, himself a physicist, starts his argument by showing that the choice of mass, length, and time as primary dimensions in physical measurement is a matter of definition and convenience and not due to the "real nature" of physical objects. In this respect the physicist is on no sounder ground than is the psychologist in his measurements. He does not object to this state of affairs, but rather insists that it must be recognized. "The habit of referring the imperfect systems of the real world to imaginary perfect 'forms' is so useful in science that one could hardly exaggerate its importance, were it not for the fact that there is to-day an undoubted danger that the physicist may forget that he is only one stage nearer 'perfection' than his colleagues who practice less exact branches of science. The difference is one of degree and not kind." Where the difficulty arises is often in the failure to recognize that "a dimensional symbol must in truth describe not a property . . . , but a process of measurement." A number of sources are cited illustrating problems in physics which show that accepted dimensional definitions are not always the most convenient ones. These thoughts are not new, nor does Scott Blair claim them to be, but his treatment puts them into a more readable form for those only somewhat familiar with the field of physics. Apparently there are physicists who lack understanding of these issues underlying the measurements they make, hence the need for argument.

What is new in the book stems from the special problem of handling dimensionally incompatible quantities, and it is here that Scott Blair looks at psycho-

logical measuring techniques as a possible way of getting out of the dilemma. Specifically, Scott Blair's practical sort of work makes it necessary for him to measure substances which lie somewhere between ideally elastic materials and viscous fluids, but classical physical measurements do not provide any way of saying that an elasticity is greater or less than a viscosity. He sees in this situation certain analogies with Gestalt psychology, with psychophysical measurements where qualitative differences may be evaluated quantitatively, with factor analytic methods of searching for meaningful psychological dimensions, and with quantitative considerations in constructing the color solid. Indeed, he describes experiments in which *jnds* of firmness as judged by *O*s are used to give physically meaningful measures. Since the term "physical property" usually implies invariable dimensions, Scott Blair prefers to call these complex properties with variable dimensions "quasi-properties." The theory of quasi-properties is not all this simple, and the interested reader is referred to the book for details. The argument for a need to cross measurement-barriers is an effective one, even if somewhat rambling in style and more suggestive than conclusive.

Anyone who takes the title of Scott Blair's book too seriously will be disappointed in the limited coverage of the topic. The audience to which its message is directed appears to be classical physicists who are intolerant of the problems of practical physicists and would be especially intolerant of use of psychological techniques of measurement in physical research. Psychologists *qua* psychologists are likely to find little except a new source for support of the claims of legitimacy of psychological measurement. In either physics or psychology the book is likely to have its greatest use as an 'outside reading' with appeal depending upon the reader's fondness for provocative scientific essays.

University of Nebraska

KATHERINE E. BAKER

L'Analyse Factorielle et ses Applications. By VARIOUS AUTHORS. Paris, Editions du Centre National de la Recherche Scientifique, 1955. Pp. 431.

During the period July 11 to July 16, 1955, an international conference of distinguished research workers in the area of factor analysis was held in Paris. The roster of participants included such well-known names as Burt, Eysenck, Guilford, Guttman, Hotelling, Ledermann, Piéron, and many others. During a period of five days of meetings there were nine sessions, in each of which two or three papers were usually read and discussed. The current volume constitutes a transcription of the proceedings.

Although some attention was given to strictly mathematical aspects of factor analysis such as those formulations presented in a paper by Hotelling concerning the "relations of the newer multivariate statistical methods to factor analysis" or in the one by Bargmann regarding a "statistical test for the stability of simple structure," the principal emphases tended to be oriented toward psychological considerations. Substantive areas were represented by such papers as the one regarding "dimensions of intellect" prepared by Guilford, the place of factor analysis in research in psychophysiology as presented by Piéron, the trends of research in spatial abilities as described by El Koussy, the use of factor analysis in the evaluation of achievement tests as described by Husen and Henrysson, the exploration of biological vari-

ability as put forth by Schreider; the application of the methods to the study of mortality as developed by Ledermann, and a recapitulation of the Princeton conference of 1951 concerning the use of factor analysis in the parsimonious description of aptitudes and personality dimensions as summarized by Thurstone (who was unable to attend the meetings). In addition certain methodological aspects of factor analysis that were psychologically oriented appeared in such papers as those of Yela, who discussed the "psychological meaning of factor analysis as a research method," of Eysenck, who critically evaluated factor analytic procedures in the light of their validity in the oretical psychology, of Guttman, who described his new "radex approach to factor analysis," or of Peel, who focused his attention upon "the factor analysis of person correlations and the use of independent determiners to identify the factors."

For the student of factor analysis whose knowledge of French is adequate enough to permit him to read the many papers presented, the volume furnishes an up-to-date and representative compilation of the work and points of view of many, if not most, of the leading research workers in the field of factor analysis. Fortunately many of the ideas in these papers, or substantial portions of the papers themselves, have been published in recent issues of American and British journals. Nevertheless this volume is an important contribution to the literature of factor analysis and a valuable reference source for the psychologist who wishes to augment his knowledge concerning the use and place of factor analysis in the study of human behavior.

University of Southern California

WILLIAM B. MICHAEL

A Twentieth-Century Bestiary. By VARIOUS AUTHORS. New York, Simon and Schuster, 1955. Pp. xi, 240.

To comparative psychology with its dearth of elementary yet readable material, of simple but scientifically respectable writing, this collection of articles taken from past issues of the *Scientific American* (and published originally under the title *First Book of Animals*) is a welcome addition. Their fitness as collateral reading for a course in comparative psychology makes one suspect that the publishers had student demand in mind. The selection and editing of the articles are designed to appeal to a literate if not necessarily scientific readership. No attempt has been made to present a balanced coverage of the field, but the choice of authors, for the most part deeply involved in and expert in their respective fields, results in a highlighting of the key problems in specific areas of research. One consequence of editorial policy that is less desirable from the viewpoint of teaching is a rather too neat presentation of some problems. Each article tends to become a little gem, pleasing to the casual reader, but lacking the ragged edges that might incite unease and raise new questions in an active mind. This premature polish exaggerates a trend already present to a disturbing degree in the professional journals. Perhaps it accounts for the observation of many a scientist-reader that he is pleased and edified to read *Scientific American* articles in all fields excepting that of his own specialty. In his own area the expert may more easily recognize when over accentuated progress and optimism are chimerical.

Of the content of the book, more than a third of the 24 articles are directly concerned with animal behavior, a few touch on the relation of behavior to ecology, two deal with evidence of evolutionary adaptations, and the remainder describe

physiological and anatomical mechanisms. The following animals are represented: birds, insects, mammals, and fish, in order of diminishing number of articles. From the point of view of teaching, the non-behavioral selections may well prove useful in giving the student some ideas of evolutionary process and some sense of the relation of behavior to the internal and specialized mechanisms of the organism as well as to its general ecological setting.

The many instructors already using *Scientific American* articles as collateral reading material will appreciate this reasonably priced selection with regret that practical considerations prohibited reproduction of all but a few of the handsome illustrations of the originals.

Brandeis University

RICHARD HELD

Reading Difficulties—Their Diagnosis and Correction. By GUY L. BOND and MILES A. TINKER. New York, Appleton-Century-Crofts, 1957. Pp. 486.

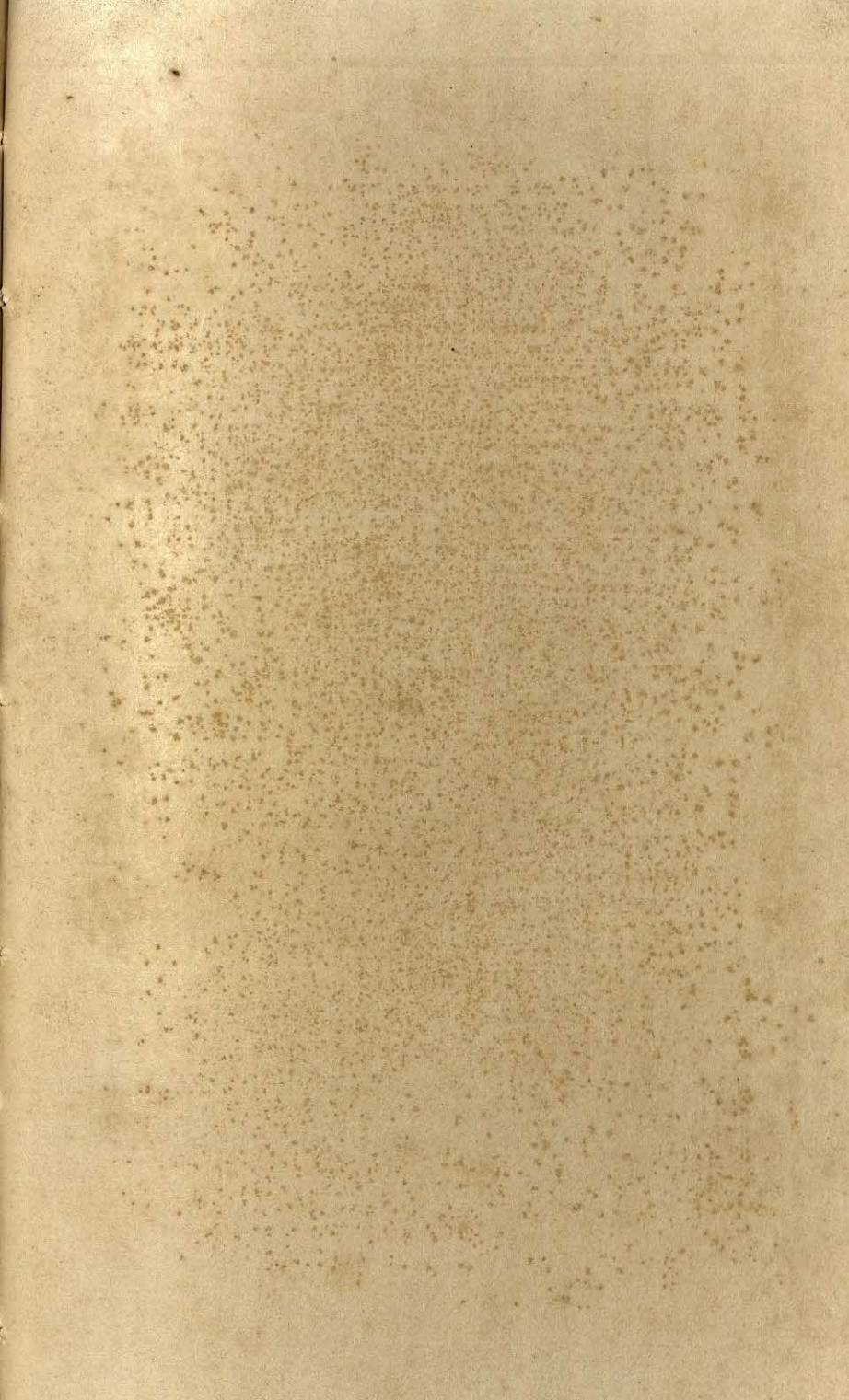
Children who have reading difficulties need expert help if they are to receive maximum benefits from education. This book is intended to serve as a basic manual for those who wish to provide help for children who have reading problems. It carefully describes diagnostic procedures and material, providing specific diagnostic approaches and techniques. It offers a variety of remedial procedures, with numerous illustrative exercises dealing with each type of remedial problem. It discusses in detail the many problems which stem from a reading disability, helping teachers understand why children with reading problems behave as they do. Case histories are scattered throughout the book, and the concluding chapter is devoted entirely to more detailed case studies, but the authors have carefully avoided the overworked, emotion-filled type of story about retarded readers. The case histories serve to re-emphasize the need for sympathetic understanding as well as better reading instruction.

The book is divided into five parts. The first part deals with the general principles of reading instruction, including the adjustment of instruction to individual differences. The second part considers the nature and causes of reading difficulties and the physical, sensory, emotional, intellectual, and educational factors involved. The third part of the book explains the principles of diagnostic procedures and the fourth part gives specific directions for the remedial treatment of each kind of reading problem. The final part discusses such special problems as improving reading tastes and adjusting to the specially handicapped child. The Appendices include a valuable list of tests and helpful bibliographies.

This book should be of great value to all teachers who are eager to help children with reading difficulties but who, too often, lack the understanding and techniques to deal with them.

Reading Clinic, St. Louis Public School

KAY WARE





Lewis M. Terman

(See page 472)



THE AMERICAN JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXX

SEPTEMBER, 1957

No. 3

MYCOTIC NATURE OF BRAIN DAMAGE IN MENTAL DEFICIENCY

By J. W. PAPEZ, and PEARL PAPEZ, Columbus State Hospital

Material for this study was obtained from 70 cases of mental deficiencies which came to postmortem during the last five years. We examined the cerebral cortex and cells of hypophysis cerebri with the phase microscope for living organisms at a magnification of 970 diameters. In each case a histological study of sections of the brain was made.

Methods. Phase method was used to show living organisms. Suspensions of small pieces of cortex were made in a drop of sterile normal saline on microslides. Cover glasses were sealed to the slides with hot Vaseline. Organisms appeared in profusion in all cases. The variety and life cycle were illustrated in reference.¹ In some stages the organisms showed characteristic motility. The elementary bodies, dot-forms and ring-forms oscillated and tumbled in Brownian movement; filaments which grew in mycotic cells moved in a sinuous manner.²

Histological methods showed the position of colonies of mycotic organisms in cytoplasm of nerve cells in most cases, and in acidophil and basophil cells of hypophysis. A special stain was used to show these mycotic parasites in nerve cells (Figs. 1 and 7), and in the hypophysis. This method removed the lipofuscin, the fat that surrounded the organisms. We examined the cortex and the thalamus by histological methods in order to confirm the presence of these mycotic organisms in these tissues; and to establish their relations to degenerative changes and lack of develop-

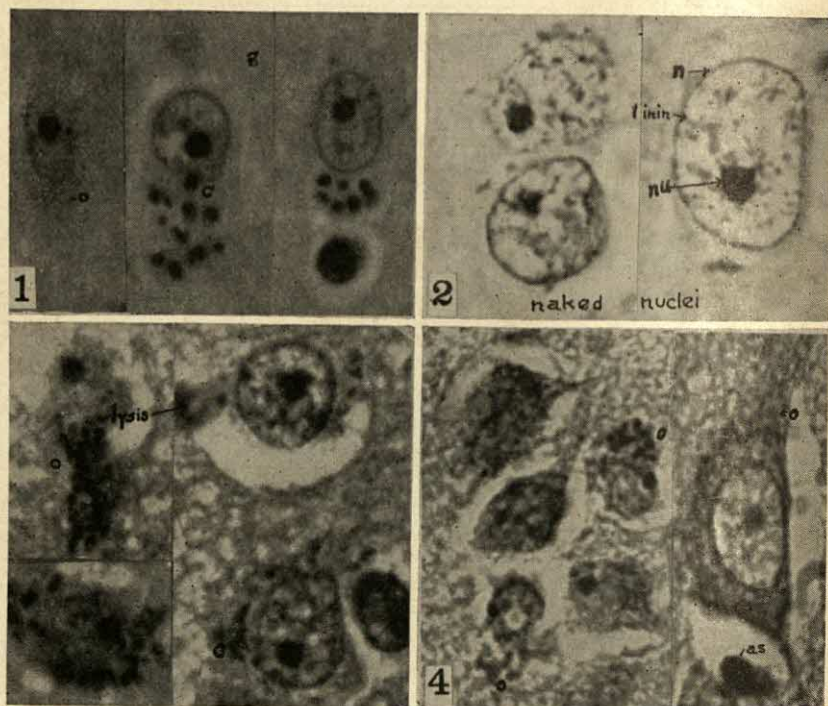
* Received for publication December 5, 1956. From the Department of Mental Hygiene and Correction and the Laboratory for Biological Research, Columbus State Hospital, Columbus, Ohio. The material, for which our thanks are due Dr. R. Gove and his successor, Dr. W. A. Butcher, superintendents, came from the Columbus State School. The autopsies were performed by Dr. A. Towbin and his successors Drs. J. M. Bloodworth and O. J. Lowy. Our thanks are also due Dr. M. R. Wedemeyer, superintendent of Columbus State Hospital where most of this work was done.

¹ J. W. Papez and B. P. Papez, The hypophysis cerebri in psychosis, *J. nerv. ment. Dis.*, 119, 1954, 326.

² J. W. Papez and J. F. Bateman, Changes in nervous tissue and a study of living organisms in mental disease, *J. nerv. ment. Dis.*, 114, 1951, 400; Form of living organisms in psychotic patients, *ibid.*, 116, 1952, 375.

ment of cytoplasm which could be observed in the nerve cells and gland cells. The following descriptions are based on these histological studies.

Cytopathological changes. Abnormal changes in nuclei and defective growth in cytoplasm were visible in the nerve cells. The glia tissue also



FIGS. 1-4. SLIDES FROM THE BRAINS OF VARIOUS PATIENTS

as = astrocytes; c = cytoplasm; g = glia; n = nucleus;
nu = nucleolus; o = organism

Fig. 1. Subject *GN*, age 22 yr.; mongolism; nerve cells with large mycotic organisms; from caudate nucleus. (x970).

Fig. 2. Subject *JM*, age 21 yr.; dementia praecox with mental deficiency. The slide shows nuclei of nerve cells devoid of cytoplasm. Biopsy from the cortex. (x1800).

Fig. 3. Subject *EH*, age 73 yr.; feeble-minded; nerve cells from striatum shedding organisms. (x970). At right, two large, almost naked nuclei from occipital cortex. (x1200).

Fig. 4. Subject *PS*, age 60 yr.; feeble-minded; dwarf nerve cells, little cytoplasm, no dendrites; on right, one larger cell with dendrites. (x970).

showed abnormal growth. These deviations were indications of metabolic and genetic disturbances seen in young and growing nervous tissues. It appeared that metabolism of cell nuclei was altered, and that the growth

of cytoplasm was retarded.³ Damaged cytoplasm was altered in a pathological direction. It showed holes, tatters, deformities and poverty of cell processes. Overgrowths of glia tissue and dysgenesias of nerve cells were associated with the presence of the mycotic organisms. During the first decade these organisms were difficult to demonstrate because of their small size or absence. The phase microscope, however showed them. In sections the minute mitochondria-like organisms did not take the stain, but their presence was inferred from minute holes in the cytoplasm and from the phase microscope studies.

Cytopathological changes in the nuclei of nerve cells were prominent.⁴ Chromatic granules were often numerous and linin threads were thick (Fig. 2). It looked as if their robust growth was accentuated by some abnormal growth-inducing substance. Nuclear sap that made up the unstained contents was increased. There was an increase in size of the nucleoli and an increase of chromatic granules on the linin threads that stained like desoxyribonucleic acid. In severe cases and in older individuals the nuclei were densely stained and deformed (Fig. 10), a change referred to as pycnosis.⁵ In other cases the nuclear sap was greatly increased and nuclei became large without cytoplasm (Fig. 2).

Mycotic changes. We have inferred the presence of some hyperergic substance that produced abnormal growth. It was possibly a mycotic auxin, acting on the protein threads of the nuclei. A purple substance interpreted as a mycotic auxin (5-hydroxytryptamine) appeared in the organisms under the phase microscope.⁶ Pycnosis of nucleus was often associated with pycnotic shrinkage of cytoplasm (Fig. 6). These changes were usually similar to those seen in brains of schizophrenic patients in whom there was better cell development. In the feeble-minded there was a lack of cytoplasmic development to the extent that the cells were small and often without dendrites or axons. They gave the impression of being immature and without function. This condition of poor development affected the pyramidal cells of Layers iii and v as well as the small granule cells of

³ Papez and Bateman, Cytological changes in nerve cells in dementia praecox, 110, 1949, 425; Papez, *ibid.*, 114, 1951, 400; Cytological changes in cells of thalamic nuclei in senile, paranoid and manic psychoses: Significance of dorsal thalamus in psychoses, *ibid.*, 112, 1950, 401. See also *J. clin. exper. Psychopathol.*, 12, 1951, 89.

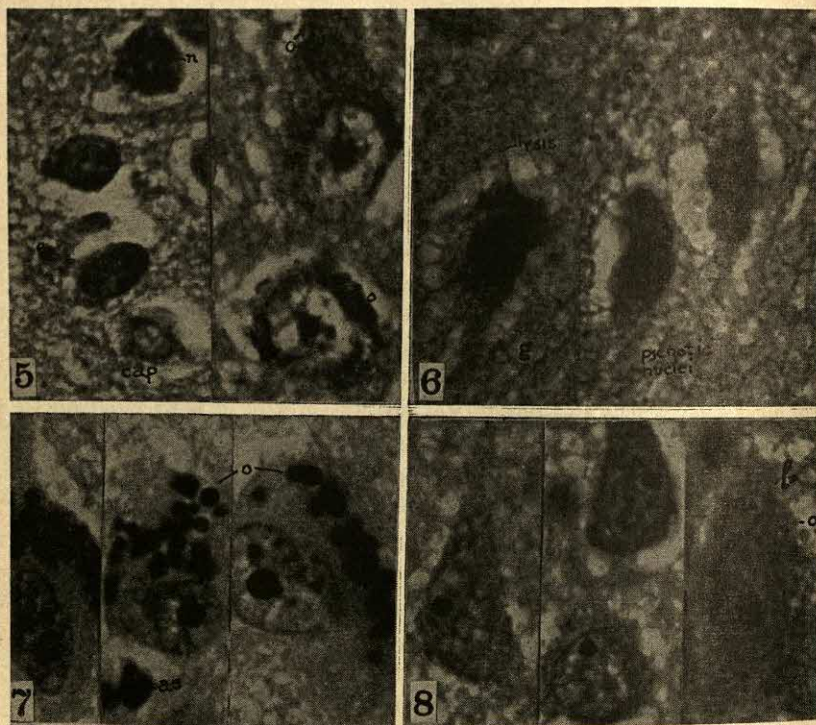
⁴ Papez and Bateman, 110, 1949, 426-431.

⁵ Papez and Bateman, 114, 1951, 408.

⁶ J. W. Papez, Lipid and mycotic nature of arteriolar sclerosis in cerebral vessels and related condition, *Trans. Amer. neurol. Ass.*, 1956, 184-186; Papez and Papez, Drops of protein in brains of hospital patients in stupor, uremia, edema, clouded and catatonic states, *J. nerv. ment. Dis.*, 124, 1956, 361-370; Arteriolar mycosis associated with chronic degenerative brain disease: A new look at sclerotic patches, *Dis. nerv. System*, 18, 1957, 2-8.

Layers ii and iv of the cortex. It seemed that specific functions of these cells were affected. (For a summary of proposed functions of granule cell layers, see Papez 1956.)⁷ Study by decades follows.

First decade. We studied 17 cases in the first decade of life. In young



FIGS. 5-8. SLIDES FROM THE BRAINS OF VARIOUS PATIENTS

as = astrocytes; cap = capillary; g = glia; l = lysis;
n = nucleus; o = organisms

Fig. 5. Subject *MM*, age 60 yr.; feeble-minded; cell nuclei without cytoplasm or processes; on right, two large cells with mycotic spores. (x970).

Fig. 6. Subject *CF*, age 22 yr.; feeble-minded; dwarf nerve cells with pycnotic nuclei, lysis of cytoplasm; fine mycotic spores at o. (x970).

Fig. 7. Subject *GN*, age 22 yr.; mongolism; nerve cells with large mycotic organisms from striatum. (x1500).

Fig. 8. Subject *RB*, age 32 yr.; damaged cortical nerve cells, cytoplasm honey-combed with fine organisms. (x970).

patients with mental deficiency the genetic changes in nerve cells stated above were more severe than in older patients. The cytoplasm in the nerve

⁷ J. W. Papez, Granule cells, their distribution in central nervous system, considered as inhibitor system, *Anat. Rec.*, 121, 1955, 411.

cells of the cortex and thalamus was destroyed in its early formative stages, before the cytoplasm had a good start (Fig. 11, left). The cell nuclei were small and lacked complete cell bodies. Their bodies lacked dendrites and also axons in most cases. As a result, the nucleus of the nerve cell had only a small, tattered rim of cytoplasm without its usual extended branches. Axons and axon sheaths were rare.

This agenesis of cytoplasm and its branches appeared in sections of the thalamus and of the cortex in young patients. Nerve cell nuclei devoid of cytoplasm were seen in brains of idiots and imbeciles who showed few or no signs of intelligence. In brains that lacked cellular layers in the cortex, indications were that the dysgenic process began before birth, or in early infancy.⁸ We have found that the brains of aments had few nerve cells which were usually grossly defective (Fig. 4). These defects were associated with rather scanty presence of mycotic elementary bodies as if the disease had consumed the cytoplasm at an early age. This condition was in some cases continued into later years. In the feeble-minded, dysgenesis of cytoplasm was partial, but still severe, affecting some but not all of the nerve cell. Cell population was sparse.

Second decade. We examined 12 cases from the second decade of life. Some brains with scant development of cytoplasm, as in the imbecile, were continued into the second decade. In the feeble-minded, those with better developed brains had better formed cell bodies with dendrites. There were more good axons, and more nerve fiber sheaths of myelin and glia tissue, but still very few axons. In the cytoplasm of the nerve cells there were large colonies of small inclusion bodies, or the cytoplasm was full of bubble-like holes where the organisms had been. They filled most of the cell body and occupied 60 to 90% of its volume. We assumed that the cells could not function normally, if at all, with such a heavy load of mycotic spores and so much of their cytoplasm destroyed. A gross damage of cytoplasm and absence of its synaptic surface were prominent features of our findings in most cases. The cell population was markedly reduced in number. Cytoplasm of small granule cells suffered most (Fig. 6). The cell layers of the cortex had, in many cases, an irregular and random arrangement lodged in a confused mass of fine, overgrown glia tissue, where few or no axons were seen (Figs. 11, 12). Others with better developed nerve cells showed cytolysis and sloughing of cytoplasm.⁹

Third decade. In the third there were five cases for study. Their nerve

⁸ C. E. Benda, *Developmental Disorders of Mentation and Cerebral Palsies*, 1952, 19, 121; R. J. A. Berry, *Brain and Mind of the Nervous System of Man*, 1928, 372, 490, 500.

⁹ Papez and Papez, 124, 1956, 361-370.

cell population was much reduced and represented largely by pycnotic, stippled nerve cell nuclei, often deformed with narrow strands of cytoplasm around them that lacked dendrites or axons. The cell surfaces were ragged, without any organisms except where there were holes or shreds on the surface to indicate they had been present. In some cases strands of cytoplasm extended out into the surrounding glia network (Fig. 6). This was a form of cell dissolution which resembled cytolysis seen in acute stages of schizophrenia.¹⁰ In some cases the mycotic spores were of large size and numerous around some of the damaged nerve cells (Fig. 7).

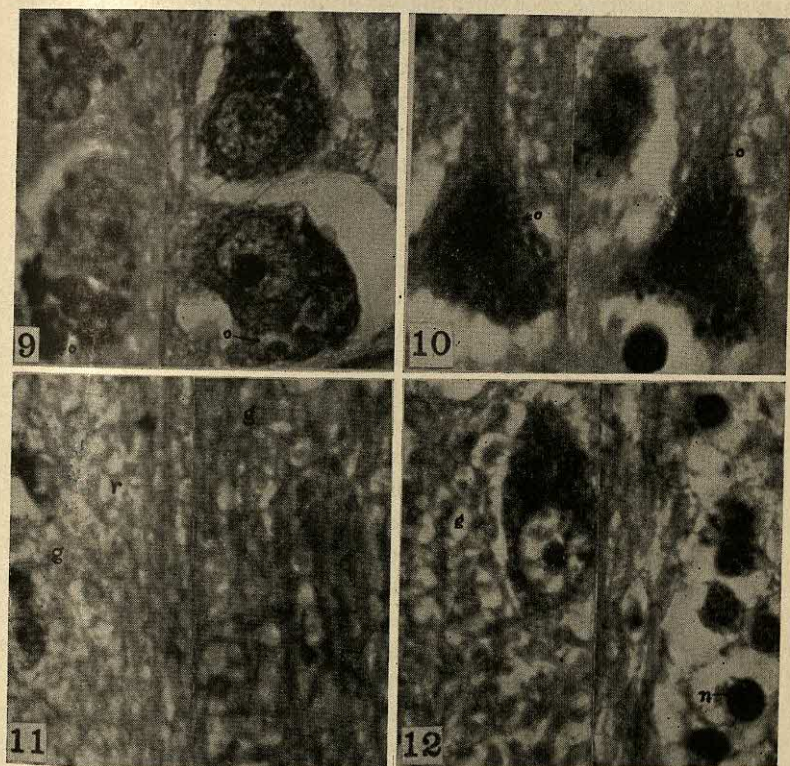
Fourth decade. In the fourth decade the eight cases we studied had in general a better cell population in layers of their cortex. The cells were larger with severely pitted cytoplasm (Fig. 8). Mycotic spores were small or rarely of medium size. Cytoplasm was blistered, shredded or full of holes. Dendrites were absent. Few axons were seen. These were defective cortices in severely retarded individuals. Their nerve cells were better formed than those seen in younger brains. Nuclear pycnosis was usual. A few giant axons were seen in radial bundles and internal capsule.

Fifth, sixth, and seventh decades. In the fifth, sixth and seventh decades there were eight cases in each for study. Cell populations were fairly good in the cortex. There was dense pycnosis of deformed nuclei which in general had scant rims of cytoplasm without dendrites or axons (Fig. 9). Naked nuclei without functional connections were the outstanding feature responsible, we think, for the mental deficiency. Organisms in the ragged cytoplasm were numerous with spores of medium size (Fig. 10). Dissolution and sloughing of cytoplasm was going on (Fig. 9 left). Cell surfaces were ragged, without contacts or dendrites. Cytoplasm was often scattered in the surrounding cell space (Fig. 3 left). Throughout these three decades, in 24 cases the histological studies presented evidence of an old original mycotic cell disease and a continuation of the same destructive process in surviving nerve cells with reduced cytoplasm.

Eighth, ninth, and tenth decades. In the eighth, ninth, and tenth decades there were three cases for study. In these decades the gliosis was intensified, coarse and confused (Figs. 11, 12). Damage around axons was severe. In a 94-yr.-old patient whose mental deficiency began at age of 16 yr. the cortex showed a damaged population of nerve cell nuclei in all layers of the cortex. The destruction of cytoplasm, cell surfaces, dendrites and axons was estimated to be 90%. There was an equivalent paucity of synaptic contacts.

¹⁰ Papez and Bateman, 110, 1949, 430; 114, 1951, 406; Papez and Papez, 124, 1956, 361-370.

Metabolic changes. Metabolic changes resembled those we described in brains of schizophrenic patients.¹¹ The presence of some abnormal nitrogenous substance such as mycotic auxin which induced abnormal growth



FIGS. 9-10. SLIDES FROM THE BRAINS OF VARIOUS PATIENTS

g = glia; l = lysis; n = nucleus; o = organism; r = radial bundles

Fig. 9. Subject *FEM*, age 58 yr.; on the right large thalamic nerve cells with large mycotic organisms. On the left, Subject *EH*, age 73 yr.; dementia praecox with mental deficiency; cells sloughing organisms. (x970).

Fig. 10. Subject *EH*, age 73 yr.; honey-combed pyramidal cells with organisms, and pycnotic nuclei; cortex. (x970).

Fig. 11. Subject *KR*, age 9 yr.; feeble-minded; remains of cells in dense glia; radial bundles with no axons. On right, Subject *RB*, age 32 yr.; radial bundles destitute of axons and nerve cells. (x970).

Fig. 12. Subject *CF*, age 22 yr.; confused dense gliosis around a lone cell with organisms. On right, Subject *RB*, age 32 yr.; dense glia tubing without axons; large fluid spaces around oligoglia nuclei. (x970).

¹¹ Papez and Bateman, 110, 1949, 430; 114, 1951, 406; 112, 1950, 401.

of brain tissue may be assumed.¹² Probably it was supplied by the organisms in the nerve cells. These changes occurred in the cortex and thalamus (Fig. 9, right) in a generalized way wherever the presence of the organisms was observed, even when the organisms were so minute as to escape recognition. They were readily seen under the phase microscope. An inadequate formation of cytoplasm and its irregularities were usual features which rendered the cells functionless.

Cerebral regions affected. Various regions of the brain were affected by dysgenesis. They were frontal, parietal, occipital, temporal, hippocampal, thalamic, tegmental, striatal, and the hypothalamus. These regions showed defective, or imperfect development of their nerve cells. In the cortex there was defective development of the cortical layers of cells and radial bundles which usually lacked axons (Fig. 11). They were often involved in an overgrowth of confused glia (Fig. 12). The distribution of these defects was general and random. In some cases the defects was continuous as if due to a local spread of disease following the course of blood vessels. This suggested that the mycotic disease came from blood vessels.¹³ In other cases the cortical defects seemed related to defects of cells in the thalamus associated with a lack of thalamocortical fibers. A diminution of nerve fibers or their oligoglia tubing was evident in the internal capsule (Fig. 12, right), corona radiata, and association fibers in the white matter of the cerebral hemispheres. The convolutions were often dwarfed or microgyric in some regions.

Thickened cortex. The thick cortex in some microcephalic brains was due to increase in glia with a lack of nerve cells or with undeveloped cytoplasm. This thickness of the cortex was an overgrowth of astroglia tissue between the remnants of diseased nerve cells (Figs. 11, 12). There was an actual reduction of nerve cells in the cortex, and a reduction of nerve fibers in the white matter partly replaced by a dense growth of oligoglia. Tubes which this glia forms around axons were usually absent or few. There was also disorganization in the intracortical fibers. The radial bundles were often undeveloped when the thalamus showed severe damage of its cells with an absence of thalamocortical radiations (Fig. 11).

Thalamic dysgenesis. Dysgenesis of cells of dorsal thalamus was severe in 26 cases of mental deficiencies that we examined.¹⁴ The cell groups which formed the thalamic nuclei were hard to recognize because they

¹² Papez and Papez, 124, 1956, 361-370.

¹³ J. W. Papez, Strains of mycotic parasites in red blood corpuscles from hospital patients, *Anat. Rec.*, 124, 1956, 466; Papez and Papez, 124, 1956, 361-370.

¹⁴ J. W. Papez, 112, 1950, 401. Bateman and Papez, *J. clin & exp. Psychopath.*, 12, 89-103.

were disorganized and partly destroyed (Fig. 9). Usually there was a random cellular distribution and not the orderly arrangement for which we looked. In some cases nerve cells were destroyed, and others were without dendrites. The outer cell membranes which provided the synaptic surfaces were often severely tattered and otherwise damaged. Cell nuclei with thin rims of cytoplasm remained. This damage involved nerve cells, nerve fibers and probably synaptic connections. It must have placed a severe limit on the patient's receptive capacities. In some instances giant axons were present in the internal capsule. They appeared to arise from some large thalamic cells. The thalamus showed the largest and oldest organisms in most cases.

Glial changes. Abnormal supporting tissue in the gray matter of cortex and thalamus was characterized by an unusual overgrowth of astroglia.¹⁵ It was evident as increase of processes of astrocyte cells. There was an overgrowth of networks of astrocyte fibers around the nerve cells, and between nerve cells and blood vessels (Fig. 12, left). All the glia tissue in the bloodbrain barrier showed this excessive growth of fibers that probably obstructed the free flow of cerebral fluid. In ten cases of epilepsy with mental deficiency a dense gliosis was a pronounced condition.¹⁶ It probably impeded the flow of cerebral fluid to the nerve cells and fiber sheaths.¹⁷

Axonal changes. The sheaths of axons were formed by oligoglia tubes which contained myelin and in its center the axon surrounded by a fluid membrane. Oligoglia overgrowth in the white matter was a feature common to all cases, and areas. The abnormal condition of oligoglia tubes is shown in Figs. 13 and 14. We examined the fiber sheaths in the radial bundles of the cortex, in the internal capsule, in the optic tracts and basis pedunculi. Many defects were observed, such as absence of glia tubing, replaced by whorls of glia fibers of a fine texture, where nerve fibers with myelin should have been seen (Fig. 11). The myelin was especially damaged and caused separation of membrane from axon.

The spaces around the axons which normally contained cholinesterase, the so-called membrane, appeared to be enlarged or disrupted (Fig. 13). The conduction of nerve impulses was probably absent, slight or grossly impaired.¹⁸ These confused networks of overgrown glia were seen in the internal capsule, in fiber bundles in thalamus and elsewhere (Fig. 12,

¹⁵ A. R. Elvidge and G. E. Reed, Biopsy studies of cerebral pathological changes in schizophrenia and manic depressive psychosis (swelling of oligodendroglis and mild hyperthropy of astrocytes), *Arch. Neurol. Psychiat.*, 40, 1938, 227. Papez and Bate-man, *J. nerv. ment. Dis.*, 114, 1951, 410.

¹⁶ J. W. Papez, Abnormal changes in central motor apparatus in epilepsy, *Anat. Rec.*, 124, 1956, 428.

¹⁷ Papez and Papez, 124, 1956, 361-370.

¹⁸ Papez and Papez, 124, 1956, 361-370.

right). Such disorders of glia tubing were not seen in normal brains, and not to the same extent in brains of schizophrenic patients. It was suspected that some toxic fluid influence, perhaps a mycotic auxin which stimulated abnormal growth, had been at work for a long time within this damaged tubing normally occupied by cholinesterase.

Variations. Variations were numerous. The abnormal changes were not as uniform as the foregoing description might imply. The age-groups

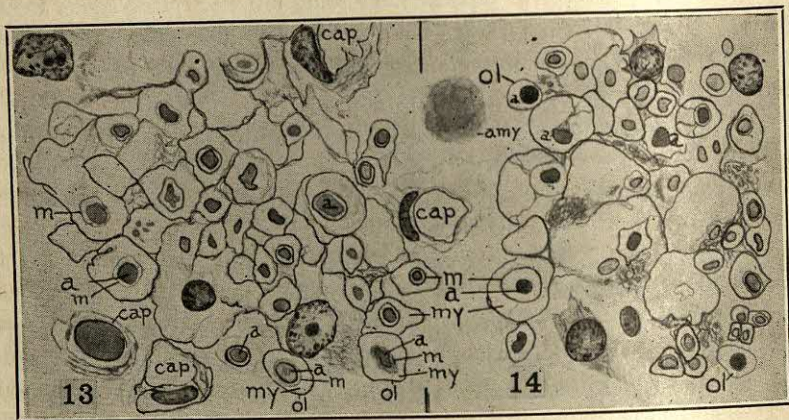


FIG. 13. CROSS SECTION OF SPINAL CORD

Subject *COM*, age 61 yr.; showing oligoglia tubes (ol) containing myelin (my). The membranes (m) around axons (a) are markedly changed. They contain an excess of cholinesterase, the conductor fluid around the axon. Myelin is degenerated. (x1200).

FIG. 14. CROSS SECTION OF TEGMENTUM IN A MUCH DETERIORATED PATIENT Subject *JH*, age 55 yr.; showing the damage membranes around axons and some giant axons. (x1200). a = axon; amy = amyloid body; cap = capillary; m = membrane; my = degenerated myelin; ol = oligoglia tube

were not uniform in pathology. In different cases there were regional variations peculiar to the case. Different areas and often different vascular fields were involved. In hydrocephalus there was ependymal disease, and also mucoid exudate infested with living organisms. In some of the cases with epilepsy the glia was extremely dense and indurated. The frequent variations were mainly the agenesis of cell cytoplasm in the young. They were responsible for most of the dysgenic defects in the cortex and thalamus. These small cells lacked dendrites, axons and neurofibrils. They occurred mainly in idiots and imbeciles when these survived the period of childhood. They were irreparable cytogenic defects as judged from the histological studies.

Blood vessels. The blood vessels were examined in most cases. In many

instances they were damaged. In others, they were enlarged and deviated by mycotic spores. In some cases there were fine granular parasites in the plasma and in the red blood corpuscles. In some vessels a great thickening (arteriosis)¹⁹ was present in the intimal lining of the vessels.²⁰ The specific nature of the granules was not determined. In some cases we recognized them as the fuzzy forms of mycotic spores and as the peculiar granular parasites in corpuscles in dementia praecox.²¹ From experiments with cultures in animals we think they might have been the same as 'blood platelets' we found in mice and guinea pigs after injections of brain suspensions taken from human biopsies.²² In another article we have described the step by step formation of mycotic (sclerotic) patches on blood vessels of the brain seen in this series of cases.²³ We have illustrated this form of red blood cell mycosis.²⁴

Hypothalamus. The hypothalamic nuclei have been examined in a few of these cases. The large cells of the supraoptic and paraventricular nuclei were found severely damaged.²⁵ They sent their fibers by way of the hypophysial stalk to the pars nervosa of the hypophysis. These fibers passed between the pituicyte cells and terminated in the homogeneous membrane which covered the outer surface of the lobules of this part of the gland.

We examined the hypothalamus grossly in cases where histological specimens were not secured. The pars optica was generally thin and undeveloped.²⁶ In some cases the vascular plexus and pars tuberalis of the gland were inconspicuous as if they had failed to develop. The tuberal region where the dorsomedial and ventromedial hypothalamic nuclei were located was in some cases a thin lamina with narrow cellular contents, and the third ventricle was dilated with a thin floor. Provisionally, we were of the opinion that the hypothalamic nuclei were probably as defective in histology as their gross appearance indicated. Neural regulation of food intake takes place in this part of the hypothalamus.²⁷

¹⁹ Benda, *op. cit.*, 329.

²⁰ Papez and Papez, 124, 1956, 361-370. N. W. Winklemann and J. L. Echel, Endarteritis in small cortical vessels in severe infections and toxemias, *Arch. neurol. & psychiat.*, 21, 1929, 863.

²¹ J. W. Papez, *Anat. Rec.*, 124, 1956, 466; Papez and Papez, *J. nerv. ment. Dis.*, 124, 1956, 361-370.

²² Papez and Bateman, Reticulo-endothelial response following injections of human brain tissue into mice and guinea pigs, *Trans. Amer. neurol. Ass.*, 1951, 219.

²³ Papez and Papez, 124, 1956, 361-370.

²⁴ Papez and Papez, 124, 1956, 361-370.

²⁵ L. O. Morgan, Alterations in the hypothalamus in mental deficiency, *Psychosomat. Med.*, 1, 1939, 496-507.

²⁶ Morgan, 496-507.

²⁷ J. R. Brobeck, Neural regulation of food intake, *Ann. N. Y. Acad. Sci.*, 63, 1955, 44-55.

Hypophysis. The pars nervosa of the hypophysis was usually filled with swarms of mycotic elementary bodies as seen under the phase microscopes.²⁸ They came from the small round cysts that occupied the membrane around the lobules of the pars nervosa. In some cases the shattering or sporulation of these cysts was observed. During this process the elementary bodies oscillated as they came out of the surfaces of the cysts. The origin of the cysts was not clear. It seemed possible that they came from the nerve cells of the hypothalamic nuclei in the manner usually ascribed to neuroendocrine secretions in this region of the hypothalamus. We came to recognize the membrane of pars nervosa as an important repository of these round purple cysts.

Dorsomedial and ventromedial nuclei. The dorsomedial and ventromedial hypothalamic nuclei occupied the region of the tuber cinereum. They were examined in 16 cases. They consisted of small and medium-sized nerve cells. The ventromedial nuclei, one on each side, were closely related to the hypophyseal portal veins that led down in pars tuberalis to the sinusoids of the pars anterior of the hypophysis. The small nerve cells also sent some axons to the pars tuberalis. The neuroendocrine secretions of these nuclei probably were carried into the sinusoids of the anterior hypophysis by this portal hypophyseal system. We had no way of showing this.

The cells of the hypophysis are known to secrete various hormones which influence the activities of other endocrine glands. Thyrotropin promotes the secretion of thyroid hormone which influences body growth. Adrenocorticotropin (ACTH) affects adrenal cortical hormones. The gonadotropins promote the functions of the testes and ovaries, and the somatotropins aid the growth factors in cartilage, bone, muscles, nervous, vascular and connective tissues. The blood vessels which leave the hypophysis carry these secretions to the other endocrine glands and directly to bodily tissues. A return system of blood vessels brings back some chemical influences to the hypothalamus and the gland. These are described in books on endocrinology.

Anterior hypophysis. Small pieces of the anterior hypophysis were examined in all the cases in saline suspensions under the phase microscopes for living organisms.²⁸ A heavy mycotic infestation was found in all cases in the cytoplasm of acidophil and basophil cells which produced the internal secretions described above. The hypophysis was studied in some of these cases which were stained by a special cresyl echt violet method to show the mycotic organisms in the cytoplasm and the amount of destroyed tissue. (The customary stains did not show these important disease agencies.) This destruction of cytotropic granules in these cells was of a

²⁸ Papez and Papez, 119, 1954, 336.

severe grade in the children who showed lack of skeletal growth, muscular dystrophy, cerebral agenesis, and some related malformations. These include cases of achondroplasia, muscular dystrophy, arthritis, myasthenia gravis, spastic paralysis, and dysgenesis of motor parts of the central nervous system.²⁹ The secretory cells of the adrenal glands were also seriously affected and often destroyed in these cases.

When the trophotropic mechanism of the hypothalamus was damaged, we assumed that the secretions of the anterior hypophysis were affected. The thyrotropic hormone, the general growth factor, and adrenotropin were probably affected by the organisms present in hypophysial and adrenal cells. It followed that the reduced trophotropic functions of the hypothalamus and endocrine glands were only a part of the dysgenic picture in the feeble-minded. It was our impression that the mycotic disease of the hypothalamus and anterior hypophysis described above were important factors which contributed to the bodily deficiencies in many cases of feeble-mindedness.

Clinical terms. In addition to mental deficiency the terms included idiocy, imbecility, feeble-mindedness, microcephaly, cerebral palsy, muscular dystrophy, myasthenia gravis, mongolism, hydrocephalus, Schilder's disease, and epilepsy.

Mental deficiency and dementia praecox. The pathology in brains of feeble-minded was similar to that in schizophrenics in many respects.³⁰ The growth reactions inside of the nuclei increased the linin and chromatic particles to produce pycnosis, etc. This showed that these genetic centers were behaving as if stimulated by some unusual growth promoting substance. In this regard mental deficiency could be looked on as a disease of the genetic material in the nuclei of nerve cells of the growing child. The destruction of cytoplasm of the nerve cells in infants and children was related to the activity of mycotic spores. They were of smaller size than those in dementia praecox. We regarded them as forerunners of the larger forms. When the nerve cell had grown up, in older children, the process of dissolution and sloughing of its cytoplasm with the mycotic spores took place. This was similar to that seen in dementia praecox, a cytolysis produced by dissolved lymphocytes.³¹ We supposed their lysin attacked the organisms in the nerve cells and caused the outer portions of the cells to loosen and break away. The ensuing loss of outer synaptic surfaces of the cells, their dendrites and axons was similar in the two diseases. The disconnected, functionless nuclei with their rims of cytoplasm occurred early in mental

²⁹ Benda, *op. cit.*, 31.

³⁰ F. J. Kallmann, S. E. Barrera, P. H. Hoch, and D. N. Kelly, The role of mental deficiency in the incidence of schizophrenia, *Amer. J. ment. Defic.*, 45, 1941, 514.

³¹ Papez and Papez, 124, 1956, 361-370.

deficiency before the nerve cells were fully formed. They occurred later and in a smaller number of cells in schizophrenia. In older, deficient patients with a higher IQ the condition resembled that in a deteriorated dementia praecox. We had the impression that in infants and older children the small, elementary, fuzzy, mycotic spores occurred first in the blood plasma before they entered the nerve cells. This may be the case also in schizophrenia. In mental deficiency the disease strikes before the brain is fully developed. In schizophrenia the brain is fully developed before the disease strikes.

SUMMARY

Seventy brains, hypophyses, and adrenal glands from mentally deficient patients were examined in suspensions under phase microscope that showed a heavy infestation with living mycotic spores. Histological sections specially stained showed these parasites in the nerve cells and glands. At least one area of the cortex and often several were examined in each case. The thalamus which sent fibers to the cortex was examined in some cases. The changes induced in the nervous tissues were manifold, destructive and dysgenetic. They produced abnormal growth activity (pynosis) in the nuclei of nerve cells, destruction and dissolution of the new-formed cytoplasm, with dysgenesis and agenesis. In general, nerve cells lacked an outer surface, dendrites and axons, and most of them appeared to have little or no functional connections. Cell population and cortical layers were severely reduced in all young cases. Correlation with clinical picture was definite in most cases. Overgrowth of astroglia cells and their fine glia fibers replaced the absent nerve cells. With the reduced number of axons, the oligoglia sheaths for them were absent and reduced to a redundant chaotic growth of oligoglia tissue that formed the white matter. Radial bundles were relatively absent in cortex. Thalamic radiations to cortex were as scant as thalamic cells, and there some giant axons occurred. The damaging effects on hypothalamus, anterior hypophysis, and endocrine glands were related to dystrophic development of skeleton, muscles, and the nervous system. The cytotropic defects on these tissues were complex and disastrous. They destroyed the secretory products of the cerebral hypophysis whose abnormal secretions affected adversely somatic tissues, producing bodily defects commonly encountered in persons with mental deficiency.

THE EFFECTS OF ASSIMILATION AND CONTRAST IN JUDGMENTS OF CLINICAL MATERIALS

By DONALD T. CAMPBELL, WILLIAM A. HUNT, and
NAN A. LEWIS, Northwestern University

The promise of a generalized psychology of judgment seems most evident in the effects of stimulus-context upon the response to a specific stimulus. Here such phenomena as anchoring and adaptation-level seem to hold both for a wide variety of simple physical stimuli and for more complex stimuli whose intensity varies in terms of learned connotations independent of any physical dimension.¹ The clinical materials used in the present study are an example of such complex stimuli. We used schizophrenic responses from a vocabulary test, and the judges were asked to scale them for the severity of disturbance exhibited.

That context should influence the judgment of such materials seems eminently reasonable. Thus the second author, in unquantified observations of the judgments of Naval psychiatrists performing in a screening unit during World War II, noted that psychiatrists with a background of private practice tended to consider examinees to be more disturbed psychiatrically than did psychiatrists with a background of state hospital experience, a finding which conforms to adaptation-level theory. Yet a previous study by Arnhoff using clinical judgments of schizophrenic test-responses failed to demonstrate predicted anchoring effects.² This outcome is so out of keeping with the general evidence in the area of judgment that it seemed desirable to reexamine the problem, using a design employing a drastic change in stimulus-context rather than relying upon anchoring. Our hypothesis was that stimuli of median value would be judged as less severe when presented in context with stimuli selected from the severe end of the continuum, and as more severe when presented in context with stimuli from the mild end of the continuum.

* Received for publication February 15, 1957. This study derives from a larger project subsidized by the Office of Naval Research under contract 7onr-450(11) with Northwestern University.

¹ For convenient reviews of this literature see J. P. Guilford, *Psychometric Methods*, 2nd ed., 1954, 302-340; and D. M. Johnson, *The Psychology of Thought and Judgment*, 1955, 280-389. For studies involving dimensions not physically measurable see J. G. Beebe-Center, The law of affective equilibrium, this JOURNAL, 41, 1929, 54-69; W. A. Hunt, Anchoring effects in judgment, this JOURNAL, 54, 1941, 395-403; and H. R. McGarvey, Anchoring effects in the absolute judgment of verbal materials, *Arch. Psychol.*, 39, 1943 (No. 281), 1-86.

² F. N. Arnhoff, Some factors influencing the unreliability of clinical judgments, *J. clin. Psychol.*, 10, 1954, 272-275.

METHODS AND PROCEDURE

Materials. The basic stimulus-materials, schizophrenic responses to items in a vocabulary-test, were drawn from Arnhoff's study.³ They had been standardized on the basis of judgments by 22 experienced clinicians. To this original pool of items were added some at both the normal and disoriented extremes. Further judgments were obtained from 49 college sophomores in a pilot run and these were used in adjusting the stimulus-values to the college population used as Ss. Some 50 additional items for the middle range were obtained by having 26 students sort them, using the technique of equal appearing intervals. These three sets of items were pooled and then divided into nine groups representing nine scale-points on a continuum ranging from no disturbance to extreme disturbance. The selection was made on the basis of consistency among our Ss' standardization-groups. A sample item for each scale-value follows:

- (1) Gamble: to take a chance; a risk
- (2) Cushion: padded item used for comfort
- (3) Fur: all of an animal's coat
- (4) Armory: used for centuries applied to a type of structure
- (5) Brim: outside diameter with a margin
- (6) Fur: trading markets and selling to the Indians
- (7) Envelope: something you put it in for them
- (8) Cushion: to sleep on a pillow of God's sheep
- (9) Stave: that's before, that's long before not happiness

The stimulus-items were presented in booklet form with five items on a page. Table I shows schematically the format of the booklets with the order of presentation of the items. Within each page the order of items was distributed by a balanced pair of 5×5 Latin squares to eliminate any correlation between position on the page and item value. Phase 1 of the experiment contained items from one extreme of the scale. Phase 2, a transitional phase in which the range of the items was gradually changed, was introduced to avoid any abrupt and obvious change in stimulus-context which might alert the S to the nature of the experiment. Phase 3 consisted of items from the opposite end of the scale used in Phase 1. There were two experimental groups, as indicated in Table I. The Low-High group started with the less disturbed, relatively normal items and shifted to those from the more disturbed end of the continuum. The High-Low group started with the more disturbed items and then shifted to the less disturbed ones. Items of Value 5 were presented on every page in each phase for both experimental groups. These items formed the common denominator for our basic comparisons, being first presented in a context of stimulus-items from one extreme of the scale, then in a transitional context, and finally in a context of stimuli of the opposite extreme. Changes in the judged value of the '5'—items could be related to changes in stimulus-context.

Our experimental design involves an assumption of the 'identity' of different stimuli scaled as equal on the continuum. Unfortunately, both the original standardization-data and the pilot run on the present study showed that the scale-values of the complex stimulus-materials we were using were highly variable, *i.e.* relatively unreliable. In classical psychophysical experiments using weights, tones, etc., the same stimulus can be presented again and again without the S's being aware of this

³ Arnhoff, Some factors influencing the unreliability of clinical judgments, *J. clin. Psychol.*, 10, 1954, 272-275.

repetition, or without his recognizing the stimulus on the basis of extraneous identifying cues. In social or non-metrical complex stimulus-materials of the type we were using, each stimulus-item has an idiosyncratic character of its own and may easily be recognized if repeated. Under such circumstances, as Hunt and Flannery have demonstrated, any judgment involved may be contaminated by learning.⁴ To avoid

TABLE I
DESIGN OF THE EXPERIMENTAL BOOKLETS

Low-High group	Page	Scale-values									Page	High-Low group
		1	2	3	4	5	6	7	8	9		
Phase 1	1	x	x	x	x	x					26	Phase 3
	2	x	x	x	x	x					25	
	3	x	x	x	x	x					24	
	4	x	x	x	x	x					23	
	5	x	x	x	x	x					22	
	6	x	x	x	x	x					21	
	7	x	x	x	x	x					20	
	8	x	x	x	x	x					19	
	9	x	x	x	x	x					18	
	10	x	x	x	x	x					17	
Phase 2	11		x	x	x	x	x				16	Phase 2
	12		x	x	x	x	x				15	
	13			x	x	x	x	x			14	
	14			x	x	x	x	x			13	
	15				x	x	x	x	x		12	
	16				x	x	x	x	x		11	
	17					x	x	x	x	x	10	
Phase 3	18					x	x	x	x	x	9	Phase 1
	19					x	x	x	x	x	8	
	20					x	x	x	x	x	7	
	21					x	x	x	x	x	6	
	22					x	x	x	x	x	5	
	23					x	x	x	x	x	4	
	24					x	x	x	x	x	3	
	25					x	x	x	x	x	2	
	26					x	x	x	x	x	1	

repeating items for any one S and nonetheless to have the over-all comparisons based upon identical items, counterbalancing of the Latin square type has been employed. Thus in Fig. 1, each specific item of Value 5 is represented at each point in the graph, and all points are comparable in that they are based on the same 20 specific items. The counterbalancing was achieved in two steps. To achieve a final Phase 1 vs. Phase 3 comparability, each experimental group was divided into two subgroups. The 10 items of Value 5 used in Phase 1 of Subgroup A were used in Phase 3 of Subgroup B, and vice versa, so that when the data of both subgroups were combined, all 20 items of Value 5 were represented in each phase. Within each subgroup, the specific pages in Phase 1 and Phase 3 then were so arranged, by a 10×10 balanced Latin square, that for each 10 Ss, every page appeared first for one, second for

⁴W. A. Hunt and Jane Flannery, Variability in the affective judgment, this JOURNAL, 51, 1938, 507-513.

another, etc., and that each page preceded and followed each other page once. Thus for the whole experiment there were 40 different booklet forms, *i.e.* 10 different page arrangements with each subgroup, and two subgroups within each of the basic experimental groups.

Instructions. The instructions to the Ss were as follows:

On the pages that follow you will be shown definitions of vocabulary words made by both normal and schizophrenic individuals. Your task is to rate each one of these

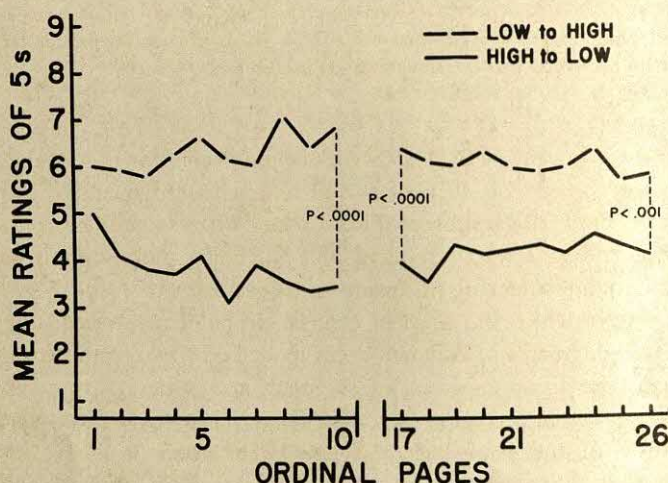


FIG. 1. SHIFTS IN MEAN RATINGS OF ITEMS OF VALUE 5

definitions according to the degree of organization and eccentricity which you think is present. You are to make your ratings on a nine-point scale which ranges from 'well-organized and normal' to 'totally disorganized and eccentric.' The category one (1) should represent the most organized and normal definitions. The category nine (9) should represent the maximal amount of disorganization and eccentricity as found in schizophrenic thinking. The intermediate categories should indicate amounts in accordance with their numerical value. You will simply write the number which you think best depicts each definition at the beginning of each statement. Be sure that you do not skip any definition, and rate each one as you come to it. It is very important that you *do not* rate definitions according to how intelligent you think the person was who made the statement. Hence, even though a definition is incorrect, if it is in no way eccentric or disorganized, it should be rated toward the low end of the scale. On the other hand, if it shows signs of disorganization or eccentricity, even though fairly accurate, it should be rated toward the high end of the scale. Try to be as discriminating as possible.

Ss were obtained from students in a course on the psychology of adjustment. Two Ss were used for each type of booklet—40 Ss for each experimental group, making a total of 80 Ss in all. The Ss worked in a group setting, taking from 20 to 40 min. for the task. Equation of the two experimental groups was achieved through randomized distribution of the booklets.

While the typical analysis of context-effects in psychophysical judgments is in terms of the point of subjective equality (PSE), and it is around this measure that

Helson has developed the quantitative theory of adaptation-level,⁵ the PSE is not suited to our materials and design. Instead we have measured context-effects by noting the shifts in judgment made to stimulus-materials of a constant value, in this case the Value 5 on our scale. While the quantitative predictions of adaptation-level theory are not readily available in this form (among other things, assumptions about range being needed in addition to predictions of mean level), this type of evaluation seems to be most generally applicable and least inferential. As it turned out, even were we to have arbitrarily taken response Category 5 for computing the PSE, the analysis would not have been feasible. The shifts in context were so great, or the adaptation-level so rigid, that the PSE in Phase 3 lay so far outside the range of the items presented that extrapolation was prohibited.

RESULTS

Shift in judgments. Fig. 1 shows mean judgments for the 5-value item. Each point is based on 40 persons and represents judgments of the same 20 words. Only the results of Phase 1 and Phase 3 will be considered here, since Phase 2 was introduced only for buffer purposes and was not given the counterbalancing treatment for equivalence.

One way in which the effect of context can be demonstrated is by comparing the judgments of common items rendered by two equivalent groups of judges, one group experiencing the items in a context of more extreme items, the other in a context of milder items. This would correspond to a comparison of the High-Low and Low-High group in Phase 1 of the present study. Fig. 1 shows the expected difference. As Phase 1 continues, the context-effect increases until at the end of Phase 1 the two groups differ significantly in the judged value of the common item, with $P < 0.0001$ by a two-tailed test.⁶ Trends within Phase 1, as tested by a t -ratio between Page 1 and Page 10, are significant at the $\frac{4}{10}\%$ level for the High-Low group, but are not significant for the Low-High group. When Mann's T -test for trend is applied,⁷ the trend for the High-Low group is significant at the 3% level, and that for the Low-High group at the 5% level.

⁵ Harry Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29.

⁶ A word should be said about the tests of significance used. Simple t -ratios have been computed comparing two points on the graph. These have not employed the correction for correlation and hence are conservative. Two source of correlation are involved but owing to the Latin square counterbalancing could be extracted only with considerable difficulty. Each point is based upon 20 separate words, and these words differ systematically in their judged values in spite of prior scaling, furnishing a correlation for all comparisons. For comparison of points on a single line there is in addition the correlation attributable to consistent individual differences among the judges.

⁷ H. B. Mann, Non-parametric tests against trend, *Econometrica*, 13, 1945, 245-259. Unless otherwise stated, the probabilities presented here are for two-tailed tests, obtained by doubling the values in Mann's Table I.

A second way to examine the effects of context would be to compare with the same group of judges the values given to an equivalent set of items seen first in one context and then in an opposite context. In the present study this would be the comparison between the judgments in Phase 1 and those in Phase 3 for each of the two experimental groups, the appropriate comparison being between the final page of Phase 1 and the final page of Phase 3 (*i.e.* between Pages 10 and 26). While the trends are in the expected direction, they do not reach clear statistical significance. For the Low-High group the difference is significant at the 2% level, but for the High-Low group the significance-level is only 23%. Trends *within* Phase 3 as measured by Mann's *T*-test are significant at the 5% level for the Low-High group, and only at the 22% level for the High-Low group.

Thus the clearly significant trends within phases are found only within the High contexts, *i.e.* in Phase 1 of the High-Low group and in Phase 3 of the Low-High group. By and large, it would seem from Fig. 1 that there is a general trend toward decreasing ratings, independent of the experimental manipulations, which might account for this asymmetry of effects.⁸

Assimilation as a context-effect in judgment. The fact that expected contrast effects appeared clearly in Phase 1 of our experiment, but were only marginal in Phase 3 after the context had been shifted for each group, led us to believe that some other antithetical tendency might be operating adversely to diminish contrast. A suggestion came from some early pilot work on this project before the Arnhoff study. Several observers volunteered the information that when working with limited ranges of stimuli at the high or low ends of a continuum, and when aware of this limitation in range, they were conscious of a tendency to develop a set toward making judgments solely within this range. Thus, if the stimulus-

⁸ This appears to be related to findings obtained by the present authors in a preliminary experiment, similar to Phase 1 of the present study, in which four items (one of Value 1, two of Value 5, and one of Value 9) were presented on each page. The judged values of the '5s' showed a significant decline from Page 1 through Page 10, although they were centered on a symmetrical set of items. The *N* in this unreported experiment was 40; the *t* between pages 1 and 10 was 2.34, $P < 0.02$. The significance of the trend had a $P < 0.03$, two-tailed, by Mann's *T*-test. Such an effect might be labeled a 'time-error,' and would make sense if interpreted in terms of a prior adaptation-level involving normal persons, which the Ss presumably bring from their relevant preexperimental experience. Hence, the 5-value items originally would appear more severe to a person previously accustomed to normal responses, but as the experimental context comes to outweigh the preexperimental adaptation-level, the '5s' would tend to appear less severe. Such an independent tendency for '5s' to decline in perceived severity would tend to augment the context effects in the High phase of the present experiment and tend to dampen out those in the Low phase, in congruence with our obtained results.

range were at the severe end, a tendency toward 'severity' would develop in their judgmental attitude, and they felt that because of this they might be judging some stimuli as more severe than they would under other conditions. Hence we felt that in our experiment the shift from one context in Phase 1 to an opposite context in Phase 3 might have made some Ss unduly aware of the new range and have produced such an 'assimilative' tendency in their judgments. Hunt had previously noted the presence of similar disturbing attitudes in an anchoring experiment involving complex stimulus-materials with non-metrical attributes.⁹

This led us to an examination of the direction of shift-effects in individual Ss. For 23 of the 40 Low-High Ss, and for 21 of the 40 High-Low Ss, the expected contrast-effects occurred, in many instances in strong degree. For a number of others, however, equally strong assimilative effects were found. For a substantial number of the Low-High Ss, the '5s' were judged higher when they were the lowest stimulus-items in the final High context than they had been where they were the highest items in the initial Low context, and vice-versa for the High-Low series. Such Ss seemingly noted the shift in average stimulus-value and changed their responses accordingly, overcorrecting, as it were, and assimilating the '5s' to the new context.

The statement that some Ss show assimilation and others show contrast can easily be confused with a naïve and unwarranted interpretive reification of chance variation. To test the validity of this apparent double-directional effect, a *t*-ratio was computed for each S, expressing the difference between the mean judgment of the '5s' in Phase 1 and Phase 3. The distribution of these *ts* for the Low-High group ranged from 4.44 in the assimilation direction to 7.04 in the contrast direction. Of the 40 Ss in this group, 9 showed *t*-ratios in the assimilative direction in excess of the 5% value of 1.734, which is significant with a *P*-value < 0.0002. Five of the *t*-values exceeded the 1% value of 2.552. In the contrast direction, eight values exceeded the 5% level (*P* < 0.0008) and five, the 1% level. For the High-Low group the range of *ts* was from 7.22 in the assimilative direction to 6.01 in contrast, with seven assimilators beyond the 5% level (*P* < 0.0004) and eight contrastors beyond 5% (*P* < 0.0008). Four exceeded the 1% level in each direction. We interpret these values as establishing the existence of both assimilation and contrast in the present population.

Differences in assimilation and contrast. In an effort to increase our understanding of the conditions giving rise to assimilative effects in Phase 3,

⁹ Hunt, 401; this JOURNAL, 54, 1941, 395-403.

we searched for possible predictors in the Ss' performance in Phase 1. It seemed possible that those Ss who had used the entire range of available ratings for the limited stimulus-range of Phase 1 were forced to redefine the categories in a contrasting direction to be able to report all the discriminations made among the items of Phase 3. To test the possible effect of such initial prodigality in the use of vocabulary with subsequent exhaustion of vocabulary, the mean of the Phase 1 ratings was examined. For the Low-High group, those Ss with higher mean ratings of all the items of Phase 1 should show greater contrast effects, and contrariwise for the High-Low group. To test this prediction every S was assigned the mean difference of Phase 1 and Phase 3 ratings of '5s' as a score on an assimilation-contrast continuum. The correlation between the Phase 1 mean rating of all items and this assimilation-contrast score was 17% for the Low-High group and 1% for the High-Low group, values which do not support the hypothesis of vocabulary-exhaustion.

A more direct test of this hypothesis might consider the tendency to use extreme category-ratings in the direction in which the shift is to occur. Thus an S who in Phase 1 of the Low-High group used a large number of Rating 9 might be forced to redefine this category to express the higher intensities appearing in Phase 3, with this redefinition generating a contrast-effect. All Ss in the Low-High group were scored on the frequency of their use of Rating 9 in Phase 1, and the High-Low group for the use of Rating 1. This measure provided correlations of 0.13 and 0.22, respectively, with the assimilation-contrast scores. Although these *rs* are in the expected direction, they do not attain significance at the 5% level. We conclude, therefore, that contrast effects are not a product of exhaustion of vocabulary.

We have suggested that assimilative effects were a function of the ambiguity of the stimulus-items. If so, those Ss who show the greatest difficulty in discriminating among the items should show the greatest tendency toward assimilation. As a measure of subjective ambiguity, *i.e.* the tendency for all items to be judged alike, we took the *SD* of all judgments made in Phase 1. A low *SD* should indicate a failure in subjective differentiation, and a high *SD* should indicate high differentiation. The *SDs* of Phase 1 correlated with the assimilation-contrast scores 0.48 for the Low-High group and 0.45 for the High-Low group, both values being significant at the 1% level.

Another measure of stimulus-ambiguity might be the accuracy of the Ss' judgments, *i.e.* their agreement with the true scale-values (from the original standardization-data) of the stimulus-items. Correlation coeffi-

cients of this sort were obtained for each *S* using his judgments of Phase 1. These were transformed into *z*-scores and correlated with the assimilation-contrast scores. Coefficients of 0.54 and 0.44, significant at the 1% level, were obtained for the Low-High and High-Low groups, respectively.

It thus appears that the *Ss* who find the stimulus-items confusing and ambiguous in Phase 1 tend to show assimilation in Phase 3. The two indices used above, however, are not completely equivalent. It is possible to have an *S* making clear-cut discriminations with a consequent large *SD*, but who is not employing the common dimension of discrimination and who consequently shows low agreement with the true scale-values. We therefore correlated the individual *SDs* and *z*-scores. For the Low-High group the correlation was 0.34, and for the High-Low group 0.04, indicating considerable independence, with each contributing some separate variance in predicting the assimilation-contrast score. When the *SDs* and *z*-scores are combined in a multiple correlation coefficient, the values are 0.70 and 0.62, respectively.

Of the indices of Phase 1, there is one which can be regarded as an unplanned composite of *SDs* and *z*-scores. This is the difference between the mean ratings given to items of Value 1 and Value 5 for the Low-High group, and the mean of the '9s' and the '5s' for the High-Low group. A narrow range of means could result from a narrow range of assigned ratings (low *SD*), or from an indiscriminate assignment of ratings to items of varying values (low *z*). The range of means of Phase 1 correlated with the assimilation-contrast scores 0.69 for the Low-High group and 0.57 for the High-Low group. These values are in close agreement with the computed multiple correlations.

Temporal weighting of experience. In his initial presentation of adaptation-level, Helson examined the temporal weighting of experience in determining adaptation-level, and found that for weight-lifting no differential weighting for primacy or recency of experience was needed within the 'specious present,' or single laboratory session.¹⁰ Following Helson's formulation, the terminal points for our two experimental groups should coincide, for at that point the items compared are being judged against the same background of over-all experience. At that point the groups differ only in order of presentation. Instead, we find the groups significantly different from each other, $P < 0.001$, at the end of the experiment (see Fig. 1).

This would seem to indicate a 'primacy' effect with the initial stimuli having a greater effect than the later stimuli. If, however, we separate the

¹⁰ Helson, 21-23.

assimilators from the contrastors and present the data for each group separately as we have done in Fig. 2, we find that the assimilators are responsible for the deviation from Helson's expected prediction, and that the contrastors are performing roughly in accordance with his prediction of equal terminal judgments.

Effects of context and shifts in context. If the mean judgments of the

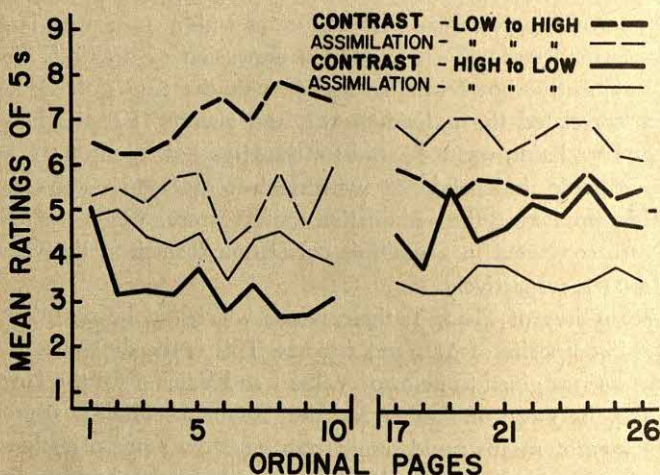


FIG. 2. SHIFTS IN MEAN RATINGS OF ITEMS OF VALUE 5 DUE TO ASSIMILATION AND CONTRAST

other stimuli are plotted (not shown here), slight trends parallel to those for the '5s' appear. The ordinal placement on the scale of the nine categories is also confirmed, although the equality of intervals leaves something to be desired. The only effect not previously noticed is a restriction in the range of mean values for the five levels of stimuli in Phase 3 as compared to Phase 1. When the low range of stimuli are judged in Phase 1 (Low-High group) the mean judgments on Page 10 range from 1.75 for the '1s' to 6.80 for the '5s.' When these same stimuli are judged at the end of Phase 3 (High-Low group), the range is from 1.40 to 4.07. For the high range of stimuli the change in range is from 3.40 for the '5s' and 8.01 for the '9s' to 5.72 and 8.22, respectively.

This restriction could have two explanations. It could reflect a contraction in the number of actual response categories employed, regardless of their accuracy, or it could reflect a lowered accuracy of discrimination without any shrinkage in range of categories used. We have examined each of these possible factors separately.

The *SD* of the judgments made by each *S* was computed for Phase 1 and Phase 3, disregarding the accuracy of the judgments or the stimuli being judged. These *SDs* were treated as scores and the effects upon them of both context and phase examined. For this purpose the design of the experiment may be treated as a 2×2 Latin square and analyzed as a cross-over design.¹¹ Table II presents the data in terms of cell and marginal means to make inspection of trends easier. The *F*-ratio for contexts was

TABLE II
AVERAGE *SDs* BY PHASE AND CONTEXT

Phase	Context		
	Low	High	
1	2.53 _{LH}	2.55 _{HL}	2.54
3	1.79 _{HL}	1.86 _{LH}	1.83
	2.21	2.16	

Marginals by experimental group: LH=2.19; HL=2.17.

0.48, $P > 0.05$, and for phases 108.11, $P < 0.01$. Thus there is no effect of context upon variance of the judgments, but a highly significant effect for phase. Phase 3 judgments are much more restricted in range than are Phase 1 judgments.

Such a drop in variability again is amenable to two interpretations. It could be a product of the shift in context with the resulting enforced redefinition of judgmental categories, or it could be an ordinal effect continuous throughout the experiment and independent of shift in context. If the latter were so, one would expect trends *within* each phase paralleling the one *between* phases, *i.e.* a continuing decline in variability. To check this we computed a measure of variability for each ordinal page of Phases 1 and 3. Because of the counterbalanced design, and to obtain sufficient instances to provide stable indices, the 40 *Ss* within each experimental group were pooled, and the *SD* of their 40×5 judgments was computed for each page. These measures showed no downward trend within either phase for either experimental group. For the Low-High group they ran from 2.66 on page 1 to 2.83 on page 10 in Phase 1, and from 2.10 on page 17 to 2.19 on page 26 in Phase 3. For the High-Low group the respective values were 2.57 and 2.72, and 1.83 and 2.13. These values confirm the phase differences found in the analysis of individual *SDs* but are larger in size because of the pooling of individuals. On the basis of this analysis we

¹¹ W. G. Cochran and G. M. Cox, *Experimental Designs*, 1950, 112-116.

interpret the phase differences as not attributable to such ordinal or temporal effects as familiarity, fatigue, boredom, etc. Interpreted as the result of the shift in context, the loss in response-variance corresponds to an effect found in anchoring studies when a remote anchor is introduced after a frame of reference has already been established. In the anchoring literature this effect has been given such labels as expansion of category width, extension of the absolute scale, shrinking in perceived stimulus-values, or loss of refinement of discrimination.

To check on a possible drop in the accuracy of discrimination, we computed two correlation coefficients for each *S*, one for Phase 1 and one for Phase 3, representing the agreement between the established scale-value of the items and the *S*'s judgments for the entire population of 50 items. These *r*s have been converted to *z*-values and used as scores of individuals for exploring the effects of phase and context. Table III presents the *r*-values corresponding to the average *z*s for the cells and margins. In the cross-over analysis of variance for the *z*s, two effects were significant. Context provided an *F*-ratio of 43.77, $P < 0.01$, with the low context materials the more discriminable; and phase provided an *F*-ratio of 17.56, $P < 0.01$, with Phase 3 having the lower *z*s.

As with the *SD*s, an examination of within-phase trends in the level of correlation was made, computing for each phase the correlation between the item values and the pooled judgments of the 40 judges. Again, there were no within-phase decremental trends. For the Low-High group the *r*s are 0.63 for Page 1 and 0.62 for Page 10 in Phase 1, and 0.37 for Page 17 and 0.39 for Page 26 in Phase 3. For the High-Low group the respective values are .43, .56, .51, .46.¹² The loss in fineness of discrimination is thus borne out by both variability of judgments and the correlation between judgments and scale-values of the items.

Since we have demonstrated two types of reaction to a shift in context, *i.e.* contrast and assimilation, the question arises as to whether or not this loss in discrimination is specific to either contrast or assimilation. The contrastors and assimilators therefore were compared for loss in *SD* and loss in *z* between Phases 1 and 3. While the two groups do differ in absolute level on both measures, the amounts of loss are approximately equal. The loss of discrimination in Phase 3 is typical of both contrastors and assimilators. In the Low-High group the mean loss in *SD* is 0.69 for the contrastors and 0.64 for the assimilators ($t = 0.20$). For the High-Low

¹² The apparent trends here are not supported by the values of the intermediate pages. Thus the set of 10 pages for Phase 3 runs 0.51, 0.47, 0.55, 0.50, 0.55, 0.51, 0.53, 0.054, 0.49, 0.46.

group the figures are 0.66 and 0.83 ($t = 0.79$). The mean loss in z for the Low-High group is 0.33 and 0.25 ($t = 0.97$). For the Low-High group there is a slight mean gain in z , because the effect of context is slightly stronger than the effect of phase. The gains are, respectively, 0.06

TABLE III
AVERAGE z 'S BY PHASE AND CONTEXT

Phase	Context		
	Low	High	
1	.654LH	.528HL	.596
3	.574HL	.454LH	.517
	.617	.493	

Marginals by experimental group: LH = .562; HL = .553.

and 0.07 ($t = 0.23$). The effects on variability and accuracy are thus no different for the assimilators and the contrastors.

DISCUSSION

Our work has confirmed that general principles of judgment generated in classical psychophysical settings involving homogeneous, uncomplicated stimuli and a simple dimension such as weight or size can be applied to more complicated materials such as schizophrenic test-responses judged on the complex dimension of degree of disturbance manifested. The dependence of the judged value of the stimulus upon the context within which it is experienced conforms to the previous findings in such areas as affective equilibrium and adaptation-level. Moreover, in line with previous studies of anchoring effects is the fact that judges who are allowed to develop stable patterns of judgment to a limited range of stimuli with a fixed number of response-categories and who are then presented with an expanded range of stimuli with no additional response-categories, show expansion of category-breadth, lessening of fine discrimination, and stretching of the judgmental scale. Demonstrating the generality of these principles with clinical materials offers the hope that they may be of practical value in understanding and controlling the complex judgmental processes of everyday living. In particular, it encourages the possibility of further refining the validity and reliability of clinical judgment, a necessary prerequisite of improved clinical prediction.

That our findings are sometimes not as clear as those appearing in classical psychophysics is not surprising when one considers the complexity of both the stimulus-materials and the judgmental dimensions we

used, as well as the Ss' unfamiliarity with them. As earlier work by Hunt has shown, in such complicated judgments, many extraneous factors such as previous experience, understanding of the task, self-instruction, etc. may play a much greater role than in the simple, laboratory situation.¹³ Any general theory of judgment must ultimately be expanded to handle such factors if it is to explain satisfactorily judgmental phenomena outside the laboratory.

On a simpler level, however, our findings suggest the need for amplification of any general theory of judgment, such as Helson's adaptation-level formulation, to handle such factors as amount of experience and order of stimulation. Current theory is relatively insensitive to these, yet in our findings they influence both the mean level of response to a common stimulus and the fineness of discrimination among stimuli. Order of experience is important, as well as total experience.

Finally, we have found a new phenomenon, assimilation. The discovery of Ss who show an assimilative 'illusion' rather than the expected 'illusion' of contrast, adds another complication to our understanding of the effects of context. Assimilation seems not to be anticipated by any of the current theories of judgment. Our work reveals some indication of predictive factors correlated with assimilation, but whether a monotonic mathematical theory of adaptation-level can be constructed to predict such assimilative effects must remain in doubt until we have further knowledge of them.

SUMMARY

Judgments of the degree of disturbance manifested in schizophrenic responses to items from a standard vocabulary-test were studied in relation to changes in the context of the items presented to the judges. In a context of responses exhibiting high disturbance, responses of medium value were judged as less severe than when they were presented in a context showing only mild disturbance. In some Ss a shift in context produced a tendency toward assimilation rather than contrast. Such assimilators showed less differentiation in their judgments in the initial context. Shift in context had the effect for both assimilators and contrastors of reducing the differentiation among the items in the new context.

¹³ Hunt, this JOURNAL, 54, 1941, 395-403.

INTERACTION OF THE EFFECTS OF CUES INVOLVED IN JUDGMENTS OF CURVATURE

By O. W. SMITH and P. C. SMITH, Cornell University

The problem of the perception of distance has been customarily discussed by enumerating the traditional cues which may serve as the basis for either apparent depth or scalable distance. These cues are customarily treated as single, isolated variables even though the observer's (*O*'s) optical array in his daily world is composed of combinations of them. At present, there is neither theory nor empirical evidence for asserting which of the traditional cues or combinations of them determine *O*'s spatial behavior when the optical array is a composite—a combination of variables.

An apparatus was so constructed as to control and to display combinations of cues and variables. The test-object was a perfect semicylinder. In preliminary observations judgments of the degree of its curvature varied with markings on the test-surface, type of viewing (monocular or binocular) as well as with the distance of observation. The effects of particular cues or variables were not additive in any simple fashion, but varied with the composition of the combinations. This study was undertaken, therefore, to test the hypothesis that cues or variables interact in their effects upon judgments of the curvature of a three-dimensionally curved object such as a cylinder.

The judgment of curvature was used as a criterion of distance because: (1) it is not relative to other distances in the field, but only to the height and depth of the object itself; (2) it is a function of change in the slant of the surface; and (3) apparent continuity of the curvature beyond the visible limits of the surface is reported when *O* has an unrestricted view of the cylinder, thus furnishing an index of normality or veridicality of judged shape.¹

APPARATUS AND PROCEDURE

A semicylinder, 7.96 cm. in radius and 16.2 cm. long, was so mounted on a central horizontal axis that the line of *O*'s regard bisected the width and height of the surface and was perpendicular to its horizontal axis (see Fig. 1). Along the inside

* Received for publication March 18, 1957. Our thanks are due Professor Albert E. Michotte for sponsoring and providing space for this research in his laboratory at the University of Louvain, Belgium and for countless observations of Professors de Montpelier and Knops and Licéensée Coen-Gelders. The study was supported in part by the Psychological Laboratory of Louvain, by the Belgian American Educational Foundation, and by Contract NONR 401(14) between Cornell University and the Office of Naval Research.

¹ A. E. Michotte and Anne Coen-Gelders, personal communications.

edges of the semicylinder, at A in Fig. 1, spring clips held removable paper jackets which smoothly fitted the exterior surface of the cylinder.

At O's near position of observation, at G (116.3 cm. from the center of the front surface of the cylinder) plain white jackets (without markings) had a mean luminance of 0.12 ft. L. over their entire surface. The means of 10 readings on an

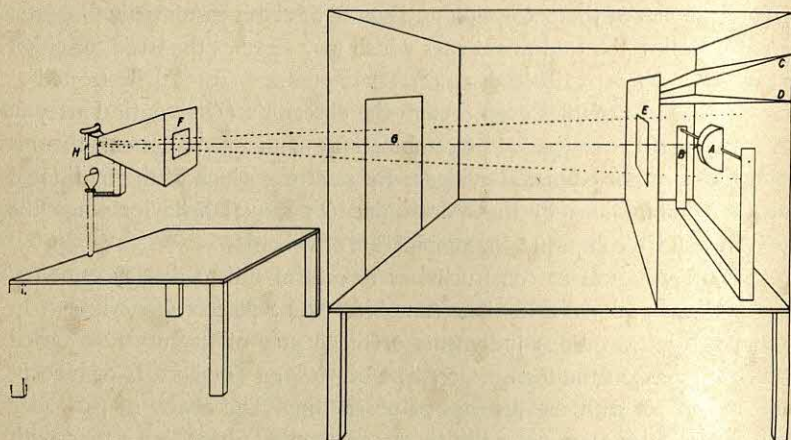


FIG. 1. LATERAL VIEW OF THE APPARATUS
(Viewer at far position)

- | | |
|-------------------|---|
| A = semicylinder | E = variable aperture |
| B = fixation spot | F = fixed aperture |
| C = shield, solid | G = position, near fixation (116.3 cm.) |
| D = shield, paper | H = position, far fixation (232.6 cm.) |

S.E.I. photometer were the same for the top, bottom, and middle of the cylinder.² The light from four overhead, 40-w. cool-light fluorescent tubes was shadowed by C and the edges diffused by papers at D. Lateral reflections were reduced by a black curtain around the test-area while papers reflected light up from the table. Uniformity of reflected light from the test-surface was tested also by unaided visual inspection. When the cylinder was covered by a plain jacket, rotated, and viewed binocularly from either the near position at G or from the far position at F (232.6 cm. from the front center of the cylinder) a flat surface was seen which resembled a window shade being raised or lowered against a neutral gray ground.

Three patterns of markings on the removable paper jackets consisted of 135, 45, and 9 black spots with an additional red fixation-spot in the center of each (see Fig. 2, 45-spot pattern). The 135-spot pattern was the same as the 45-spot pattern except for additional spots, similarly distributed. The 9-spot pattern was composed of the same spots in the same locations as marked by a cross (+) in Fig. 2. The use of different shapes minimized the effects of form- and size-constancy as cues to

² Salford Electrical Instruments, Ltd., *The SEI Exposure Photometer* (manual distributed by Ilford, Ltd., Ilford, London).

distance-differences. Spots were added at either end to minimize sharp brightness differences at the upper and lower edges. The two dotted lines in Fig. 2 enclose the area calculated to be visible from H in Fig. 1.

O saw the width of the cylinder plus areas above and below in which was seen the gray back-drop—9 cm. behind the horizontal axis of the cylinder. The lateral

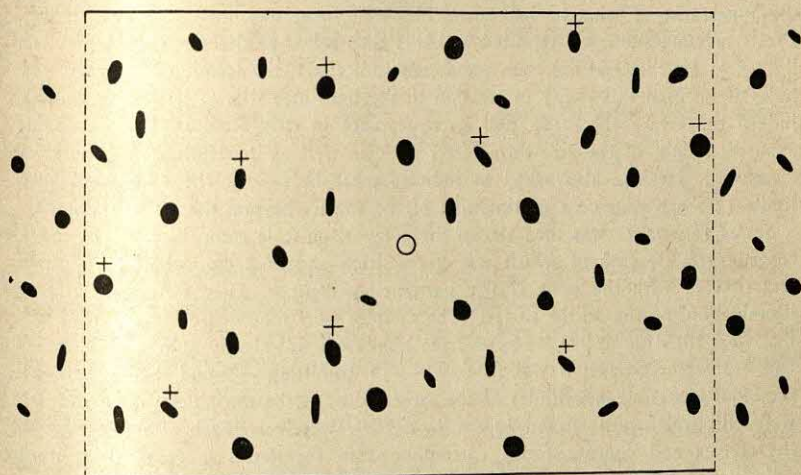


FIG. 2. PATTERN USED ON PAPER JACKET OF SEMICYLINDER
(Spots, $N = 45$)

- Limits of jacket within which spots were visible; frequency count was within these limits.
+ Spots selected from 45-spot pattern for the 9-spot pattern.
O Central fixation spot.

edges of the cylinder were occluded at both viewing positions. The cylinder subtended vertical visual angles of $7^{\circ} 20'$ and $3^{\circ} 48'$ at the near and far positions. At near and far positions 0.51 cm. and 0.26 cm. of the circumference of the cylinder at both the top and bottom were occluded since the lines of regard of O were tangent to the surface at these points.

Aperture F (16.5×14 cm. and 30 cm. from O's eye) limited O's field of view to the immediate area of Aperture E (11.5×14.3 cm. and 13.5×16.2 cm. at near and far distances respectively) and excluded any view of floor or table. The flared end of the viewing box contained a flap which E operated manually in controlling O's viewing time.

Before O was permitted to view a test-pattern, his line of regard (cyclopean eye for binocular vision and right eye for monocular vision) was brought into perfect alignment with the fixation-point. In monocular vision a removable ground-glass screen shielded the left eye but provided it with good illumination.³ Usually O was not aware that his vision was monocular.

³M. W. Horowitz, An analysis of the superiority of binocular over monocular visual acuity, *J. exp. Psychol.*, 39, 1949, 581-596.

After his visual acuity had been tested,⁴ *O* was given instructions standard for the experiment in the language appropriate to him. Among our *O*s, 60% received instructions in French, 30% in English, and 10% in Flemish. *O* was told that a shutter would open revealing a gray field pierced by a window. In the center would be a light area with a red spot in the center. He was to fixate first on the red spot and then to inspect immediately the area around it. After the view he was to describe the appearance of what he had seen.⁵ Shape or form was not mentioned, to avoid effects of suggestion. He reported and was then asked to describe the shape of the light area (if that had not been previously indicated), the location of the spots (or holes if he used this term) in relation to the light area, the materials of which he judged the area to be made, and its appearance in cross-section. He indicated the apparent height of the area on a meter stick as well as its depth, if curvature was mentioned. He was also asked to indicate what he saw at the upper and lower limits (a sharp edge or a continuation of the surface beyond the visible limits).

Scaled reports. *O* was then shown the comparison-scale comprising 1 flat, and 13 curved surfaces, each of which was 8 cm. high and 16.2 cm. wide. Each surface was covered with the same clearly patterned wallpaper. They were seen in room illumination, 60 lux on the surface. The samples were mounted on four curved racks that were tilted toward *O*, 174 cm. from him. Each was clearly numbered: -3, -2, and -1 were concave; 0 was flat; +1-+12 (omitting +9 and +11) were convex. Plus 8 was a semicylinder of the same radius as the test-object; +10 and +12 were elliptical forms, 10 and 12 cm. in depth respectively. Each of the numbers of the convex and concave forms corresponded to the depth in centimeters of the cylindrical sections.

O was told that these were samples of different forms representing a scale by which he could judge the degree of curvature or lack of curvature of what he might see in the apparatus. The range from concave through flat to convex forms was pointed out together with their numbers. He was to look again in the apparatus, then to give the number of the form (after looking again at the scale) which resembled most closely the form of the display. After his observation and judgment, he was then asked to select any other scale-values which were equally as good as his first.

Experimental design and observers. The variables in this study were: (1) ocularity—monocular vs. binocular vision; (2) number of spots on the surface of the semicylinder—3 patterns consisting of 135, 45, and 9 black spots of mixed round and elliptical shapes; and (3) position—two distances of observation at 116.3 cm. and 232.6 cm. Eleven *O*s were assigned at random to each of 12 cells in a $2 \times 3 \times 2$ factorial design, $N = 132$. The *O*s were students at the University of Louvain, Belgium, and other members of the university community. Each had 20/30 vision or better in both eyes.

RESULTS

(1) *Free report: (a) Surface.* A solid substance was reported by 131 of the 132 *O*s. Presumably they saw a surface. One *O* reported a fog with

⁴ Each eye was tested at 5 m. by a Snellen chart with an illumination of 60 lux.

⁵ The exposure-time was approximately 5 sec. and exceeded the time necessary for the development of maximal stereoscopic acuity for central vision. See K. N. Ogle, *Researches in Binocular Vision*, 1950, 116.

spots in it. Differences in the kind of surface (soft or hard substance), or in location of the spots showed no significant relationship to conditions of observation or to judged curvature.

(b) *Size*. The test-area was judged to be smaller than its real size ($p <$

TABLE I
FREQUENCIES OF FREE REPORTS OF CURVATURE (CONCAVE OR CONVEX)

Number of spots	Monocular			Binocular			Both		
	near	far	total	near	far	total	near	far	total
9	4	0	4	3	0	3	7	0	7
45	0	0	0	8	3	11	8	3	11
135	0	1	1	8	2	10	8	3	11
Total	4	1	5	19	5	24	23	6	29

0.01) under monocular observations. Size-judgments under other conditions were not significantly different from true size.

(c) *Curvature*. The first reports of the *O*s were a check against possible effects of suggestion of curvature upon the reports using the scale for judging. Since the distribution of results (see Table I) shows differences which are in general agreement with the more sensitive scaled judgments, they will not be analyzed in detail. Curvature was noted by only 29 *O*s (22%). Of these, 24 observed binocularly, with 19 at the near position.

(d) *Edge*. Thirteen of the *O*s reported surface to continue beyond the upper and lower limits. They all observed binocularly, 10 from the near position.

(2) *Scaled reports*. *O*'s score was the algebraic mean of the matches which he judged to be equally good. These scores were submitted to an analysis of variance (see Tables II and III, and Fig. 2). The main effects and double interactions were all significant at 5% level or better, even when the levels of significance are doubled, as suggested by Lindquist, to compensate for non-homogeneity of variance in monocular and binocular conditions.⁶

The most consistent differences in judged curvature were between monocular and binocular conditions. Although interactions with number of spots and distances were significant, greater curvature was seen with binocular vision with all numbers at both distances. The other main effects should be interpreted with regard to their interactions with other variables. Fig. 3 shows that an increase in the number of spots (particularly from 9 to 45) is accompanied by judgments of greater curvature, but the increase

⁶ E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1953, 81-83.

is greater under binocular conditions, and at the far position. Similarly, a decrease in distance of observation increases judged curvature, but to a much greater extent for binocular vision and for the 9-spot pattern.

The averaging of positive and negative values of the scaled reports

TABLE II

MEAN OF SCALED JUDGMENTS

(All matches judged equally good by each o were averaged. Mean of their means— $N=11$ per cell—are reported.)

Number of spots	Monocular			Binocular			Both		
	near	far	combined	near	far	combined	near	far	combined
9	+ .77	— .18	+ .30	+5.41	— .09	+2.66	+3.09	—0.14	+1.48
45	+ .41	+1.16	+ .79	+6.00	+4.23	+5.12	+3.20	+2.70	+2.95
135	+1.59	+ .77	+1.18	+6.95	+4.95	+5.95	+4.27	+2.86	+3.57
Combined	+ .92	+ .58	+ .76	+6.12	+3.03	+4.58	+3.52	+1.81	+2.67

TABLE III

ANALYSIS OF VARIANCE OF MEANS OF SCALED JUDGMENTS

(Score per O is the mean of all matches judged equally good.)

Source	df.	Mean Square	F	p
Ocularity	1	481.56	112.92	<.001
Number	2	50.61	11.79	<.001
Distance	1	96.77	22.56	<.001
Oc $\times$ N	2	18.22	4.25	<.025
Oc $\times$ D	1	62.76	14.62	<.001
N $\times$ D	2	21.25	4.97	<.01
Oc $\times$ N $\times$ D	2	7.73	1.80	
Within	120	4.29		

might have caused the absence of reported curvature in monocular conditions. Two checks for this were made. (1) The first scaled report, whether concave or convex, was given a positive sign. These values were submitted to an analysis of variance which is not reported here. F -values were slightly higher than for the reported analysis, but the shapes of curves were closely similar, and the triple interaction remained insignificant. (2) The frequency of reports of concave, convex, and flat or any combination thereof, for each of the first two choices was also tabulated (see Table IV). The frequency of reports in which both choices were curved changes with conditions of stimulation in the same way as does mention of curvature in the free report. After a second presentation, however, 34 additional Os who observed binocularly, and 41 who observed monocularly, judged the surface to be curved in both their scale judgments.

Accuracy of judgment as indicated by two choices of convex was much greater in binocular than in monocular vision, and somewhat greater at near than at far position and with 45 or 135 rather than 9 spots.

The direction of any perceived curvature was much less evident under

monocular conditions. This ambiguity of the perceived spatial form is indicated by a change in judgment between first and second choices (see last row of Table IV). The range of possibilities acceptable in the case

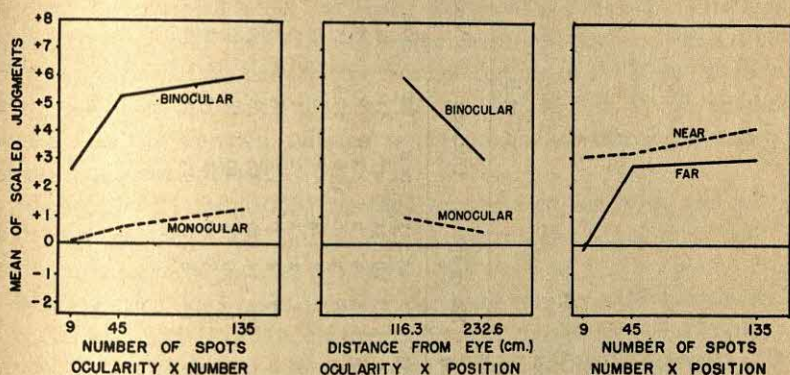


FIG. 3. RELATIONSHIP INVOLVED IN SIGNIFICANT INTERACTIONS

of monocular view is greater than with binocular vision, except in the case of 9 spots at far position.

DISCUSSION

(1) *Comparison of the reports.* Differences among conditions were in the same direction for free as for scaled reports. Curvature was reported much more frequently, however, after the second view followed by the scaled reports. Despite the instructions, intended to present concave, convex, and flat as equally acceptable alternatives the shift in number of reports of curves may have been due to the suggestion implied in the presentation of the scale of judgment. O's attention was also directed to the form of the display before the second report. Opportunities for the operation of both suggestion and attention were presumably available under all conditions so that the differences in the resultant judgments are to be attributed to other factors.

(2) *Monocular-binocular differences:* (a) *Apparent curvature.* The mean judgments under monocular conditions were all near zero as contrasted with much higher means under binocular conditions (see Table II). Except for one very small reversal, the difference appeared in all of the monocular-binocular comparisons. The effect of ocularity is the largest in the experiment, and appears in the free reports as well as in the matches. With monocular vision only 5 Os (8%) reported curvature on first view.

On second view 20 Os (30%) considered 'flat' as at least one alternative match on a scale with 14 choices available. Even when the first matches are averaged without regard to sign (a procedure which includes all possible effects of suggestion, giving maximal estimates of curvature), the largest mean obtained under monocular conditions is 2.09 (for 135 spots, near), only one half scale value above the equivalent algebraic mean. The changes in the vertical density (the size and spacing of the spots) as they were viewed at increasing slants at the top and bottom of the surface were not, in this situation, sufficient to produce any substantial impression of curvature, monocularly.

Binocularly, 24 Os (36%) reported curvature on first view, and only 8 (12%) considered 'flat' as a possible alternative. The largest mean without regard to sign (again, for 135 spots, near) was 6.82.

Much greater, and more frequent curvature was reported binocularly (see Tables I and II). The differences between monocular and binocular vision confirm findings of long standing.⁷ The traditional explanation is that the disparity effects due to stereopsis are a much more sensitive cue for differences in depth than are the so-called monocular cues. Without amplification, however, the traditional explanation cannot account for the extreme flattening found in this study under monocular conditions, nor for the interactions of variables.

The flatness under monocular conditions may be related to the tendency to see the display in the plane of the aperture, since the difference in depth between the edges of the aperture and the cylinder was not as evident as in binocular observation. The decrease in judged height of the light area, from 14.8 cm. to 10.2 cm. (corresponding to a possible reduction in apparent distance) is consistent with this hypothesis. The aperture revealed no lateral edges of the cylinder, and only a small, relatively undifferentiated area of gray backdrop above and below it.

It is also possible that there is an interaction between acuity for individual spots and the effects of density under monocular conditions. Thus, the low level of illumination used might reduce the acuity for individual spots below some critical level for the operation of monocular density effects, even though each spot was individually discriminable and reported by Os. This explanation requires experimental verification.

(b) *Ineffectiveness of monocular density.* Increases in density on the surface were ineffective in increasing substantially either frequency or degree of reported curvature at either distance under monocular conditions (see Tables II, III, and IV). This failure cannot be attributed to the use

⁷ M. D. Vernon, *A Further Study of Visual Perception*, 1952, 94.

of an inadequate range of densities. The visual angles between spots were computed at the center and the upper limit of the cylinder for each number of spots at both the positions of view (see Table V). The ranges of the retinal densities for different parts of the cylinder and for different numbers of spots are large at both near and far distances of *O*. The magnitudes of neither the angles nor their ratios are in the same order as the magnitudes

TABLE V
AVERAGE ANGLES SUBTENDED ON RETINA BY VERTICAL
SPACES BETWEEN CENTERS OF SPOTS

Position	No. of spots	Location of spots		Ratio	Reported curvatures	
		center	limit	limit/center	mon.	bin.
Near	9	66'48"	9'18"	.14	10%	68%
	45	14'36"	18"	.02	5%	75%
	135	4'36"	6"	.02	20%	87%
Far	9	33'24"	4'48"	.14	-2%	-1%
	45	7'24"	18"	.04	14%	53%
	135	2'24"	0"	.00	10%	62%

of the slight judgments of curvature obtained. There is no indication here that density changes were operating as a differentially effective monocular cue.

(3) *Ambiguity*. Two changes in density of stimulation in the optical array were a function of the increase in distance from *O* of the upper and lower portions of the cylinder. Vertical density increased as a positively accelerated function of the change of slant of the surfaces for both concave and convex forms. In this way, both types of forms generated qualitatively equivalent density changes. At the same time, each form generated unique patterns on the retina since horizontal angles subtended by spots and their interstices decreased in size as a function of their vertical distance from the center of a convex form and conversely for a concave surface.

Despite this retinal difference, direction of curvature was reported ambiguously under monocular conditions, both for the same and for different *O*s. Only 24 (36%) were certain of convexity; 7 others were certain of concavity, another 7 of flatness, and 15 *O*s (24%) considered concave or convex matches equally good (see Table IV). Binocularly, on the other hand, 54 *S*s (82%) were certain of convexity, with other reports correspondingly reduced.

(4) *Interactions*. The most important result of this experiment is the significance of the first order interactions. Both number and position show significant interactions with ocularity and with each other. Judgments of

curvature in this situation are not, therefore, simple functions of the degree of binocular disparity.⁸ Rather than discuss these interactions separately, we shall propose a more general formulation which we hope will also explain the obtained interactions.

We propose that the judgment of curvature may be judged to be dependent upon a series of just noticeable differences (j.n.d.) in depth of adjacent portions of the surface,⁹ and also upon a series of j.n.d. in elevation of these adjacent portions.¹⁰ The problem is first to define and describe the variables of stimulation which affect judgments of differences in depth and elevation. The discussion will be limited to problems involving binocular vision since the monocular judgments were inconsistent with respect to both the existence and direction of curvature. A motionless head and motionless field of view are of course assumed.

(5) *Discriminations required for perception of curvature.* Judgments of curvature are hypothesized to be a function of the rates of change of a series of ratios between j.n.d. in depth and j.n.d. in elevation included within adjacent portions of a surface (j.n.d. in elevation refer to the vertical axis of the retinal image). Each ratio can be expressed as the number of j.n.d. in depth per j.n.d. in elevation. This ratio approaches zero at the center of the surface, and, under optimal viewing conditions, becomes very large at the upper and lower limits. The factors which limit the numbers of j.n.d. which form these ratios differ according to the part of the curved surface involved. That these limiting factors exist is important only because they determine the rate of change of any one series of ratios, and consequently the degree of curvature that is judged under any one set of conditions. The rates of change of the various ratios determine whether a reported surface will be judged as flat, slanted, or curved. They also determine whether a surface is judged to be continuous or discontinuous.¹¹

The most important factor is the presence of j.n.d. in depth since these

⁸ Disparities within the cylinder were in all cases well above the minimal of 5" cited by Ogle (*op. cit.*, 116) for effective disparity. The maximal disparities on the cylinder were 12' 58" and 3' 13" for the near and far positions respectively, when the red dot in the center of the cylinder and a point just visible to both eyes beyond the aperture edge at the top of the cylinder compose the disparity pairs. (These angles are slightly larger than the comparable angles between red dot and a point at the center top of the cylinder.)

⁹ Julian Hochberg and Olin Smith, Landing strip markings and the "expansion pattern": a program, preliminary analysis, and apparatus, *Perceptual and Motor Skills*, 2, 1955, 81-92.

¹⁰ The j.n.d. is used here only as a convenient unit for discussion. A quantal unit might be preferable in some ways.

¹¹ For a discussion of this problem see J. J. Gibson, *Perception of the Visual World*, 1950, 77-144.

limit the judged depth through which curvature can take place. Reduction in the number of j.n.d. of depth will result in a reduction of judged curvature. We propose the hypotheses that: (1) any disparity of the angles subtended by two points to the eyes which is above threshold for depth serves primarily to locate one point behind the other and will be relatively ineffective for determining the extent of that separation; (2) the apparent differences in depth between two adjacent points will not be proportional to the extent to which the angular disparity is in excess of that necessary for one j.n.d. in depth, since disparities equivalent to fractional j.n.d. either are not discriminated or are ineffective in establishing changes in slant of a surface;¹² and (3) the introduction of a third point intermediate between two adjacent points separated in depth by two or more j.n.d. will increase judged depth.¹³ This last effect will increase judged depth only when there are fewer than a yet-to-be-determined optimal number of j.n.d. sufficient to specify a fully curved surface. Increase in the number of available j.n.d. beyond that number merely furnishes redundant stimulation.

For disparity to function as a variable for veridical judgments of differences in depth, two or more components must be discriminably different. Secondly, it is presumed that their lateral separation on the retinas must equal or exceed the limen for vernier acuity on at least one retina. Thirdly, the points must be sufficiently large, or sufficiently in contrast with their background, that neither the points nor the diffusion circles surrounding them fall within Panum's fusional areas on the retina.¹⁴ Previous findings on retinal disparity as a cue for judged differences in depth of two points must be revised before they will be applicable to judged differences in depth of adjacent portions of surfaces.

Near the center of the front surface of the cylinder, the change in depth is very small relative to a given change in distance along the

¹² In the case where the criterion is a scale-model, fractional j.n.d. in depth which are not judged can result only in losses in judged degree of curvature. It can be hypothesized that not only a minimal number of j.n.d. in depth is necessary for the full curvature to be judged, but that their distribution over the surface must be optimal. The above assumption was partially validated in preliminary observations when depth was judged at both the top and bottom of the cylinder when the center horizontal area on the test-jacket was unmarked. In the unmarked area the surface was always judged to be flat by both naive and sophisticated *O* ($N = 8$).

¹³ This point was validated by other observations. When seven circles, $\frac{1}{4}$ in. in diameter, were uniformly spaced on the vertical center axis of the cylinder, and observation was binocular from near position, very little curvature was judged. When the number of circles was doubled the judged degree of curvature increased, but not proportionally.

¹⁴ Increases or decreases in the size of spots even within the clearly visible range were found, in other observations, to increase or decrease reported curvature sharply. (For effects of size, See Ogle, Ch. 7.)

circumference (altitude). The factor determining the minimal j.n.d. of depth between two points on the surface, one located at center front and another higher or lower in the field would obviously be the limen for binocular disparity. Below that distance, the effective ratios would be zero, and the surface would appear to be flat. Near the top or bottom limits of the cylinder, depth changes very rapidly in relation to changes in vertical distance. The retinal angles between spots are reduced. The extent of this reduction may be estimated by the extent of the reduction of angles between spots at the higher frequencies (see ratios for 135 spots, Table V). Vernier acuity, however, requires that spots be relatively large and far apart so that there is a basis for an adequate number of j.n.d. of depth at these same limits where change is most rapid. In conflict with this requirement, spots must be sufficiently close for an adequate number of j.n.d. of depth at the same limits. (Vertical stripes would satisfy both of these requirements without conflict.)

The j.n.d. for elevation are probably much less dependent upon discriminable detail in the visual field since height on the retina may be discriminated as such. At the upper and lower limits of the cylinder, discrimination of change in elevation on the surface must involve a discrimination of the location of the line of demarcation between the cylinder and its background. If the top and bottom edges of the cylinder blend with the background, the last unit of change in elevation of the cylinder falls below one j.n.d. of elevation. Since the judged surface then ceases to change direction, it is reported as approximately hyperbolic rather than circular in cross-section. (Note that 30 out of 33 binocular Os at far position did not report continuity of surface beyond the visible limits.)

(6) *Number of spots.* We are assuming that veridical judgments of curvature would have been obtained under optimal conditions of stimulation. Since none of the obtained means approximated the value of eight, the task is to account for the loss. Under optimal conditions of stimulation the detail of the test-area would have been sufficiently large, and its reflectance and contrast of detail would have been distributed so that each adjacent component would be discriminably different from some other component. Under these conditions the rate of change of the ratios of the two types of j.n.d. would be optimally effective for judgments of curvature. The degree of the loss in perceived depth is then a function of the sensitivity of the technique of measurement and of the magnitude of the change or changes from the optimal conditions.

At near position, 135 spots, there were approximately 4,522 pairs of points for which there were measurable angles of disparity within one half of the cylinder. Yet, this number was inadequate for optimal judg-

ments of curvature based on disparity. One factor contributing to the inadequacy of disparity is that disparity loses its effectiveness with increase in lateral separation of pairs of disparate points.¹⁵

The differences between optimal and experimental conditions appeared for all patterns, but were greater for the smaller numbers of spots. For the 9-spot pattern there were 18 disparate pairs, a near-threshold number, and for the 45-spot pattern, 445 measurable angles of disparity for each half of the surface. In each case, the lateral separation of some of the pairs was so great as to decrease their potential effectiveness for disparity. The proportional loss in each case due to the relatively decreased efficiency of spots with greater lateral separation was greater for 9 than for 45 and 135 spots.

(7) *Distance of view.* If we assume for the moment that all other things are equal, and that the j.n.d. for depth and for elevation remain constant in terms of visual angle for both near and far positions, certain statements can be made. A j.n.d. in depth corresponds to a physical difference in distance between two points which must be approximately four times as great at far as at near position. Consequently, at far position, there will always be a flattening in the central frontal region, and a corresponding reduction in reported curvature.

(8) *Implications.* This hypothesis suggests that differences in rates of change of the ratios between j.n.d. in depth and elevation define differences between surfaces judged as flat, slanted, curved, or having abrupt changes, such as the junction of wall and floor. For example, in the case of a cylindrical surface, the ratios increase at a positively accelerated rate from the center to the limits of the cylinder. In the case of a flat surface such as a playing field, the ratios of j.n.d. in depth to j.n.d. in retinal elevation remain relatively constant within the ranges in which detail is discriminable. For a ground surface slanting upwards from the feet of *O* the ratios would be lower than for a flat surface. For surfaces slanting downward from the feet of *O* the ratios would be higher. A sharp change in ratios would mark the joint where a wall rests on a floor. Ratios and changes of ratios necessary and sufficient for reports of various slants at different densities remain to be determined experimentally.

SUMMARY AND CONCLUSIONS

Judgments of curvature were hypothesized to be the result of interactions of variables of stimulation in fields of view for both monocular and

¹⁵ Ogle, 48-64.

binocular vision. The judgment of degree of curvature was taken as a criterion of perceived distance. The interaction of number of spots (three levels) upon the surface of the semicylinder with the presence or absence of binocular disparity was studied at two distances of observation. Form transformation and linear perspective cues were minimized (except in the statistical sense) as were brightness differences, interposition, and other monocular effects. The following conclusions are based on free reports of what was seen as well as on judgments of degrees of curvature on a standard scale of curvature: (1) Greater curvature is reported in binocular than in monocular vision. (2) An increase in number of surface elements (spots) is accompanied by reports of greater curvature at both distances of view. This difference was very small under monocular conditions. (3) Greater curvature was reported at the nearer distances of view. (4) Monocular reports of curvature were slight, ambiguous in direction, and insensitive to differences in number of surface elements present. (5) Interactions of conditions in determining perceived distance were characteristic of judgments of curvature in this situation. (6) An hypothesis was proposed to account for the obtained interactions.

'EMPTY-FIELD' MYOPIA AND VISIBILITY OF DISTANT OBJECTS AT HIGH ALTITUDES

By R. H. BROWN, Naval Research Laboratory

Pilots and observers in planes at altitudes of 35,000 ft. or higher frequently report difficulties in the visual detection of other aircraft.¹ Among the factors listed as being contributory to these difficulties are anoxia, glare, windscreen obscuration, low contrast and small apparent size of other planes at desirable ranges of detection, the narrow field of vision for objects near the threshold of detectability, difficulty of systematically scanning the sky, and nearsightedness induced by an empty field.² The purpose of the present study is to determine the extent to which a collimated reticle corrects for empty-field nearsightedness and aids in visual detection.

A series of studies by Whiteside and his co-workers has recently emphasized the importance of empty-field nearsightedness.³ The absence of any detail in a clear blue sky, uniformly overcast sky, haze, or fog, makes it difficult for the eyes to focus at infinity. When there is no detail, the point at which the eyes focus constantly changes. In the case of a farsighted person, the average focus may be beyond optical infinity. The optically normal eye becomes on the average nearsighted when viewing an empty visual field. The nearsighted individual becomes even more markedly nearsighted under empty-field conditions.

In considering the implications of empty-field nearsightedness or myopia for further research and operational application, the magnitude and consistency of the effect are of major importance. Initially, Whiteside and Campbell measured photographically the change in curvature of the anterior surface of the lens during accommodation. In an empty field with a brightness of 200 ft.-L., the mean accommodative power for 5 Os was 0.49 diopter with a range from 0.26 to 0.68 diopter.

In a later study, Whiteside and Gronow measured the minimal visual angle for recognition of a star-shaped pattern of small dots seen at optical infinity. Of the 63

* Received for publication March 18, 1957.

¹ G. O. Emerson, Armed Forces-NRC Vision Committee, Minutes and Proceedings of the Thirty-first Meeting, Dayton, Ohio, November 1952; S. C. McLaughlin, Jr., Armed Forces-NRC Vision Committee, Minutes and Proceedings of the Thirty-fifth Meeting, Toronto, Canada, November 1954.

² T. C. D. Whiteside, Vision at high altitude, *Flying Personnel Research Committee*, Report No. 910, November 1954; M. J. Koomen, Comments on air-to-air visibility at high altitude, *NRL Memo Report No. 343*, August 1954; Ross A. McFarland, *Human Factors in Air Transport Design*, 1946, 224, 433-460.

³ T. C. D. Whiteside and F. W. Campbell, Accommodation of the human eye in an empty visual field, *Flying Personnel Research Committee*, Report No. 821, March 1953; T. C. D. Whiteside, Vision in an empty visual field: Choice of a collimated pattern, *ibid.*, Report No. 854, October 1953; T. C. D. Whiteside and D. G. C. Gronow, Vision in an empty visual field: The effect of empty field myopia upon the minimum visual angle, *ibid.*, Report No. 872, March 1954.

Ss who participated, 54 were pilots and 9 were either navigators or air gunners. In a completely empty visual field, the Ss on the average recognized the pattern of black dots against the uniform bright background when each dot subtended a visual angle of $0.82'$. When a collimated fixation-pattern was superimposed, the mean, minimal, visual angle for recognition was $0.43'$. Whiteside and Gronow also express their results in terms of the ratio between the minimal, visual angle for a distant object and the increase in size necessary for it to be visible in spite of out-of-focus blurring introduced by involuntary accommodation during search of an empty, visual field. For the 63 Ss, the mean ratio was 2.21 with a standard error of 0.82 and a range of 1.11 to 4.00.

Although the effect is small in these experiments, it is consistent. Evidence from other experiments agrees with the results of Whiteside's studies. Luckiesh and Moss,⁴ Reese and Fry,⁵ and O'Brien,⁶ found that the accommodation of the normal eye in a blank visual field of moderate or high brightness is myopic by 0.75, 1.0, and 0.94 diopters respectively.

Out-of-focus blurring produced by myopia of this order exceeds the depth of the field of the human eye. The pupillar diameter is about 3 mm. under the bright, daylight conditions of flight at high altitude. Campbell has found maximal acuity with a pupil 3 mm. in diameter for a refractive condition within 0.3 diopter of optimum.⁷ Thus, the eye's depth of field would not prevent deterioration of visual acuity if empty-field myopia exceeds 0.3 diopter. Even with a pupil 2 mm. in diameter, Campbell found, 1 diopter of myopia is sufficient to decrease visual acuity by a factor of two. Considered in terms of these results, the correction of empty-field myopia might be expected to double the range of target-detection from planes at high altitudes.

PROBLEM

Various applications of the results of Whiteside's researches have been suggested. These include provision of a collimated reticular pattern by special goggles or reflection from the mirror or windscreen of a gunsight, mounting of an accommodative stimulus on the wing tip of appropriate planes, placing collimating lenses in front of critical flight-instruments, and the use of negative lenses on pilots to compensate for myopia. Selection from among these, for use in field-tests is, however, hazardous because of two significant characteristics of research on empty-field myopia: (a) Converging stimuli have been carefully avoided by placing a fogging lens in front of one eye, while measuring the myopia for the other. Although this procedure yields the most precise information on the effect of

⁴ M. Luckiesh and F. K. Moss, Functional adaptation to near vision, *J. exp. Psychol.*, 26, 1940, 352-356.

⁵ E. E. Reese and G. A. Fry, The effect of fogging lenses on accommodation, *Amer. J. Optom. & Arch. Amer. Acad. Optom.*, 18, 1941, 9-16.

⁶ B. O'Brien, A study of night myopia, *Wright Air Development Center, Tech. Report 53-206*, May 1953.

⁷ Campbell, The depth of field of the human eye, *Flying Personnel Research Committee, Report No. 868*, March 1954.

an empty field on accommodation, the results cannot be applied immediately in the practical situation where the pilot is using binocular vision with ample cues for near vision available;⁸ (b) Targets have been presented in known positions. Information is lacking, however, on the gain from correction of empty-field myopia in a situation requiring systematic scanning of the field.

The problem of the present experiments is to test the effectiveness of a collimated reticle in improving the detection of small targets. The method employs binocular vision, cues available for near vision, and a variable position of target in the otherwise empty illuminated visual field.

METHOD

The method was selected to simulate search for other aircraft against a bright, uniform background. Fourteen Ss identified the position of a target, a dot seen at optical infinity. The Ss used binocular vision. The cues for near vision were such as a plane's cockpit and windscreen might provide. The target was systematically varied in size and in position within the otherwise uniformly illuminated field. Under these conditions, the minimal size of target for detection was measured for eight Ss in one experiment when no reticle was present and when each of three representative reticles was seen at optical infinity. In a second experiment, the performance of six Ss with the best of the three reticles previously used was compared with none.

Subjects. Fourteen enlisted men were assigned from Bainbridge Naval Training Center to serve as Ss. Of the eight Ss used in Experiment I, a medical examination indicated that six were qualified for admission to flight training. Two would be disqualified: S4 on the basis of astigmatism; and S6 because of the lack of 20/20 vision with correction. All the Ss (six in number) in Experiment II were found visually qualified. Tests of color vision and heterophoria disqualified none of the Ss for admission to flight training.⁹

Apparatus. The equipment provided for binocular viewing of a target against a uniform bright background. The cues for near vision that a pilot has in his cockpit were simulated approximately by a truncated cone directly in front of S. Thus, S looked through a wide circular tunnel and saw a target which could appear at optical infinity anywhere within the bright center, approximately 17° of visual angle in diameter. The farther edge of the tunnel, which contrasted with the bright background, approximated the distance of the windscreen from the pilot's eyes.

Fig. 1 is a schematic diagram of the equipment. A chin rest served to keep S's head and eyes in the same position for the detection of each target. A biconvex lens, 10 in. in diameter and of 20-in. focal length, made it possible to present targets at optical infinity within the laboratory. The truncated cone, in front of the lens, was so sloped that its sides appeared parallel and that its outer opening, 13½ in. from the base, appeared at an apparent distance of 40 in. when S looked through the lens.

Targets were prepared on photographic plates.¹⁰ Each plate could be placed in a

⁸ G. V. Ball, Twilight myopia, *Trans. Inter. Opt. Cong.*, 1951, 92-103.

⁹ The writer wishes to express his appreciation to Captain E. P. Irons (MC) USN, who examined the eyes of the subjects at the Anacostia Naval Air Station.

¹⁰ Mr. David S. Hawkins and Mr. Leonidas Green of the Photographic Laboratory prepared the target and reticle plates.

holder at the focal point of the lens and rotated through the plane perpendicular to S's line of sight. In preparing the 8 × 10-in. plates, contact prints were made from holes drilled with beveled edges in metal plates. These contact prints in turn were photographed and reduced in size on extremely high contrast orthochromatic anti-halation plates. The result of this photographic process was a dot, which appeared

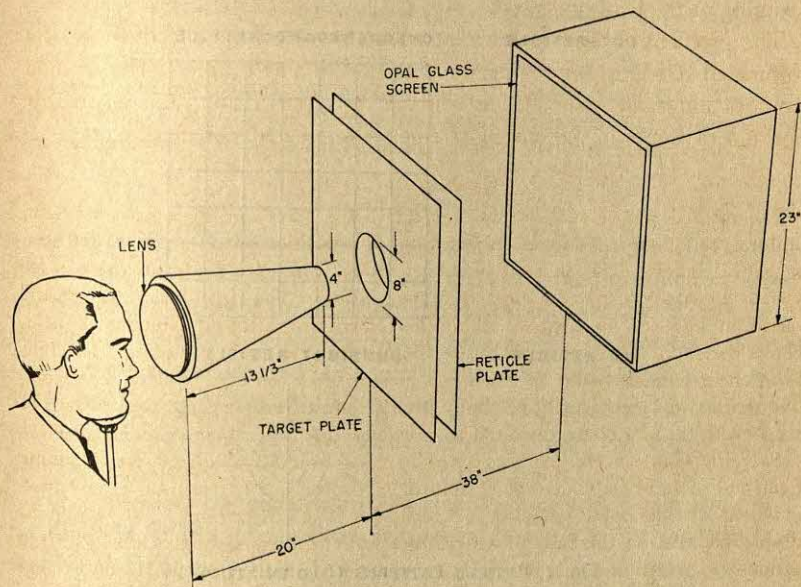


FIG. 1. SCHEMATIC DIAGRAM OF EQUIPMENT

uniformly black and circular on a clear glass plate when magnified 30 times under a microscope.

Two dots of uniform size were photographed on each plate. In Experiment I, one dot appeared $8\frac{1}{2}^\circ$ from the center of S's visual field and the other $4\frac{1}{4}^\circ$ from the center. The size of dots on the set of 14 plates decreased in uniform logarithmic steps from one subtending a visual angle of $34'$ in diameter to one subtending only $0.22'$ angle. The position of the inner dot was systematically varied with respect to the outer dot so that it might appear between the outer dot and the center of the field, on the opposite side of center, or on either side of center. In Experiment II, each dot was $1\frac{1}{2}^\circ$ from the center of the field. The two dots were separated by $43'$ of visual angle at the eye.

Each of four reticle patterns could be superimposed by means of a plate inserted $\frac{1}{2}$ in. behind the target plate. These patterns are shown in Fig. 2.

Behind the reticle plate holder was a large opal glass screen illuminated by six 20-w. daylight fluorescent lamps. At the distance of 38 in. from the reticle plate holder, the screen appeared out of focus and simulated a uniform cloud layer with a brightness of 416 millilamberts.

Procedure: (a) Experiment I. The threshold for detection of the position of the

target was measured for each reticle pattern. During a single session a given *S* made his detections for only one reticular condition. The design of a balanced Latin square was so used that each condition preceded or followed every other condition equally often. This design was selected to balance out any effect a session one day might have on another session on a later day.¹¹

Within a session, *S* reported the the position of each dot he saw, beginning with

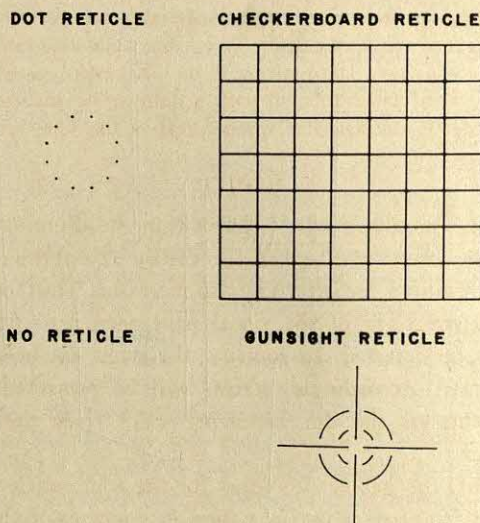


FIG. 2. RETICLE PATTERNS USED IN STUDY

the largest and progressing to the smallest. Each plate with the two dots of the same size was presented in two positions selected at random from the four possibilities of the outer dot appearing in the upper, right, left, or lower part of the visual field. After *S* had had four sessions with the four reticular conditions, this procedure was repeated but with the positions omitted during the earlier sessions.

At the conclusion of the experiment, the correct detections of the position of the dots were tabulated by size, reticular condition, and position in the illuminated visual field for each *S*. The threshold of detection was taken as the value intermediary between the smallest dot which could and the largest dot which could not be seen.

(b) *Experiment II.* In Experiment II intensive measurements were made with the best reticle pattern of the preceding experiment and with the dots of the target only $1\frac{1}{2}^\circ$ from the center of the field instead of $4\frac{1}{4}^\circ$ and $8\frac{1}{2}^\circ$. On the basis of two initial sessions, five plates were selected for each *S* to yield an anticipated 100% detection on the plate with the largest targets and 0% detection on the plate with the smallest.

In the course of ten 40-min. experimental sessions, each *S* made 10 reports on

¹¹ A. L. Edwards, Balanced Latin-square designs in psychological research, this JOURNAL, 64, 1951, 598-603.

target position, if he saw the dots, when they appeared above, below, to the right, or to the left of the reticle. *S* was allowed 10 sec. to report the position of one dot or two dots during a single observation. The plate used and the position in which the target appeared were selected at random with the restriction that each plate be used twice in each position during a session.

The dot-reticle was present in half the sessions and absent in the other half. For 3 *Ss*, the order of reticle (R) and no reticle (NR) was as follows: R, NR, NR, R, R, NR, NR, R, R, NR. For the other 3 *Ss*, the order was reversed.

At the conclusion of the sessions, the correct detections were tabulated by dot-size for each *S*, reticle condition, and position in the illuminated visual field. The number in each cell varied from 20 to 0 with a decrease in diameter of target. The threshold-diameter was calculated by interpolating at the 50% level.

RESULTS

Experiment I. Position of the target within the illuminated visual field did not affect its detection except that as distance from the center increased the size of dot required for detection also increased. Thus, on the average a dot in the extreme right of the visual field needed to subtend 1.74' of visual angle to be detected. In contrast, the inner dot needed to be only 1.25' in diameter. Accordingly, results will be presented for the more precise judgments of the dot appearing $4\frac{1}{4}^{\circ}$ from the center of the illuminated field.

Fig. 3 presents the geometric means for the four reticle conditions and for the 8 *Ss*. It is apparent that variation in diameter of the smallest perceptible dot due to reticle differences is slight and insignificant in comparison with that due to differences between *Ss*. In general, *Ss* benefited the most from the reticle with the circle of small dots which could be seen clearly only when accommodating to infinity. At best however for an individual *S*, this reticle yielded only a doubling of the visual acuity present with no reticle. In contrast, the differences between *Ss* is striking and uniform regardless of reticle-condition.

The results apparent in Fig. 3 are supported by a statistical analysis. The method of triple classification for the same individuals was used for an analysis of variance, since all the *Ss* performed all conditions of the experiment.¹² Differences between *Ss* constituted a highly significant source of variation, but differences due to reticle or to position were insignificant.

Experiment II. In view of the insignificant effect the reticle had on detection in Experiment I, it seemed desirable to make an intensive series of measurements with the target appearing in closer proximity to the reticle. Accordingly, in Experiment II, the dots were presented $1\frac{1}{2}^{\circ}$

¹² Quin McNemar, *Psychological Statistics* (2nd ed.), 1955, 311-337.

from the reticle, 68' of visual angle in diameter. It was assumed that the closer proximity of the targets should permit the reticle to have the effect observed in the earlier experiments of Whiteside and others where the target appeared in a fixed position.

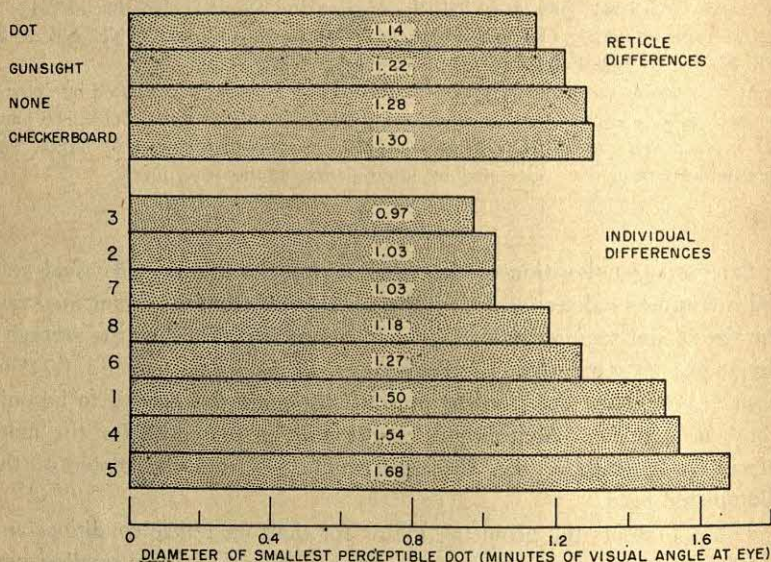


FIG. 3. GEOMETRIC MEANS FOR TARGET DOT APPEARING $4\frac{1}{4}^{\circ}$ FROM CENTER OF FIELD

TABLE I

ANALYSIS OF VARIANCE FOR DETECTIONS OF DOTS APPEARING $4\frac{1}{2}^{\circ}$ FROM CENTER OF FIELD

Source of variation	Sum of squares	df.	Variance of estimate
Reticle (R)	0.1150	3	.03833
Position (P)	0.0595	3	.01983
Ss	0.9180	7	.13114†
R × P	0.1968	9	.02187
R × S	0.8255	21	.03931*
P × S	0.5937	21	.02837
R × P × S	1.2217	63	.01939

* Significant at 5% level. † Significant at 1% level.

With the two dots appearing above, below, or on either side of the reticle, as in the first experiment, the threshold diameter for detection, on the average, was 0.561' of visual angle. Fig. 4 presents the geometric means for the two reticle-conditions, six Ss, and four positions. Again, despite

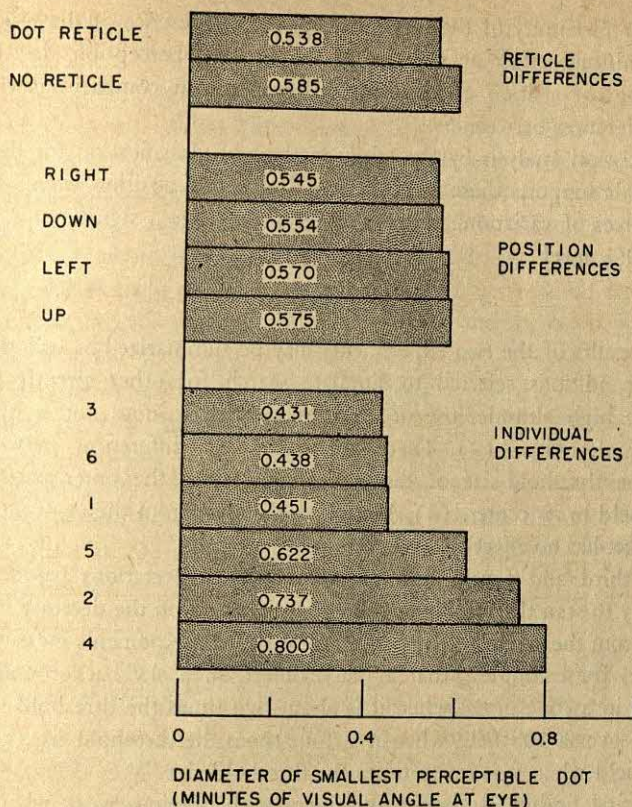


FIG. 4. GEOMETRIC MEANS FOR TARGET DOT APPEARING $1\frac{1}{2}^{\circ}$ FROM CENTER OF FIELD

TABLE II
ANALYSIS OF VARIANCE FOR DETECTIONS OF DOTS APPEARING $1\frac{1}{2}^{\circ}$
FROM CENTER OF FIELD

Source of variance	Sum of squares	df.	Variance estimate
Reticle (R)	0.0157	1	0.0157
Position (P)	0.0042	3	0.0014
Ss	0.5875	5	0.1175†
R × P	0.0231	3	0.0077*
R × S	0.0113	5	0.0023
P × S	0.0236	15	0.0016
R × P × S	0.0331	15	0.0022

* Significant at 5% level. † Significant at 1/10% level.

the closer proximity of the target to the reticle, presence of the reticle has an insignificant effect on the size of the smallest perceptible dot. Differences due to position are likewise insignificant in comparison with the large differences between Ss.

A statistical analysis by the triple method of classification for the same individuals supports these comparisons. Reticle and position are not significant sources of variation, whereas subjects constitute a significant source at the 1/10% level.

DISCUSSION

The results of the two experiments may be summarized as follows: (1) Under conditions selected to simulate search for other aircraft from a plane at high altitude, use of a collimated reticle does not significantly improve detection. (2) There are significant differences between Ss. (3) The threshold-size of target decreases from the outer part of the visual field to its center. (4) At a given distance from the center, position of a target has no effect on its detection.

The third and fourth findings agree with expectations based on the tendency to scan the middle of the visual field and on the decrease in visual acuity from the center to the periphery of the eye. Koomen's measurements indicate, for example, "that a black object on a sky background is detectable in a cone of 3° when it is about two times the threshold size, detectable over a 10° field when it is four times the threshold size, and over a 26° field when it is seven times the threshold size."¹³

The first finding is also in agreement with previous work when differences in experimental conditions are considered. Thus, Whiteside and Gronow found that use of a collimated pattern decreases the minimal visual angle for detection of a star-shaped pattern of small dots by a factor of 2.21. Requirements of binocular vision and scanning for a target which could appear in different parts of the visual field attenuate this factor to a value of 1.12 (Experiment I) and 1.09 (Experiment II). Data of the present report do not indicate unequivocally the critical factor present in air to air search which would make correction for empty-field myopia ineffective. The appearance of the target in different parts of the visual field and consequent scanning behavior by S would, however, seem to warrant further study. It may be concluded that use of a collimated reticle does not increase the range of detection effectively under experimental conditions simulating visual search for other aircraft at high altitudes.

¹³ Koomen, 3-4.

The individual differences present seem to be associated with the refractive condition of the eyes as shown by the medical examination, with the alertness and interest of Ss, and with their method of systematic scanning of the visual field. These factors have been variously reported as important in air-to-air search. For the thresholds and refractions of Experiment II, significant correlations support the belief that some degree of far-sightedness is advantageous in detecting small, distant targets.¹⁴ For six Ss, all of whom passed the visual examination for admission to flight training, the Pearson product-moment coefficient was -0.912 between the size of their target's threshold without a reticle and the correction in plus diopters required by them for 20/20 vision during cycloplegic acceptance. Thus, roughly 80% of the variance in the size of the target's threshold for this small selected sample could be attributed to slight variations in the refractive conditions of their eyes. This result suggests the possibility of selecting good high altitude observers upon the basis of refractions. We may conclude that information is needed on the relation between the refractive condition of the eye and the visual detection range for observers flying at high altitudes.

SUMMARY

A series of studies by Whiteside and his co-workers has shown during the past three years that the near sightedness induced by the enemy visual field at high altitude may account in part for difficulties in detecting planes. The feasibility of correcting for this condition has been investigated with essentially negative results in terms of design of equipment such as gun-sight reticles. Experiments indicate the possibility of improvement in air-to-air search by personnel selection.

¹⁴ J. L. Fuelling (MC) USN, Comments on Experiment 1, Personal correspondence of 17 April 1956 from U. S. Naval School of Aviation Medicine, Pensacola, Florida.

INTER-EYE DIFFERENCES IN COLOR-PERCEPTION

By ROBERT W. BURNHAM, Rochester, New York

This report concerns a method for determining the difference in color-sensitivity of the two eyes of an observer (O), and equations are presented to show how color-appearances to one eye may be predicted from a knowledge of what the other eye sees. The basic method dates back at least to Hering, but interest has usually centered on differences in sensitivity induced by grossly different conditions of adaptation of the two eyes.¹ An attempt will be made here to deal specifically with intrinsic differences in color-sensitivity when both eyes are adapted to the same illumination. The data were incidental to the results of two studies previously reported from this laboratory that concerned the effects of chromatic adaptation on color-appearance.² A binocular septum-technique was so used in those studies that a test-color could be presented to one eye adapted to a standard illuminant while a variable matching-color could be presented to the other eye adapted to the same or a different illuminant. Intrinsic inter-eye differences were completely eliminated in those studies, as in a prior study by Hunt,³ by permitting the matches made under the standard illuminant to serve as the reference-colors against which matches made with the same eye adapted to the other illuminants were compared. The difference in color-specification between the test-color presented to one eye and the matching-color seen by the other, with both eyes adapted to the same illuminant, serves as a direct measure of the difference in color-sensitivity of the two eyes and as a basis for predicting inter-eye differences in the appearance of other colors.

The usable data from the first study consisted of the results for 6 O s who matched 22 test-colors with adaptation to CIE Illuminant C but with extensive opportunity for local adaptation to the test- and matching-colors during the matching-process. The usable data from the second study comprised the results for 4 O s who matched 12 test-colors with adaptation to the same illuminant, with controlled and very limited opportunity for local adaptation while matching. Three O s were com-

* Received for publication November 12, 1956. From the Color Technology Division of the Eastman Kodak Company.

¹ Ewald Hering, Die Untersuchung einseitiger Störungen des Farbenseins mittels binocularer Farbgleichungen, *Arch. Ophthalm.*, 36, 1890, 1-23; J. F. Schouten and L. S. Ornstein, Measurements on direct and indirect adaptation by means of a binocular method, *J. opt. Soc. Amer.*, 29, 1939, 168-182.

² R. W. Burnham, R. M. Evans, and S. M. Newhall, Influence on color perception of adaptation to illumination, *J. opt. Soc. Amer.*, 42, 1952, 597-605; Prediction of color appearance with different adaptation illuminations, *J. opt. Soc. Amer.*, 47, 1957, 35-42.

³ R. W. G. Hunt, The effects of daylight and tungsten light-adaptation on color perception, *J. opt. Soc. Amer.*, 40, 1950, 362-371.

mon to both studies hence their results for the two studies can be compared for differences due to local adaptation.

Procedure. The procedure consisted of relating the mean specification of 10 matches for each test-color (left eye) to the known specification for each test-color (right eye). Specifications of test- and matching-colors were derived from spectrophotometric calibrations of the two colorimeters that were used to produce them.⁴ Since the colorimeters comprised virtually identical optical parts and were calibrated at the same time, it is reasonably certain that there were no differences due to calibration-bias. Significant calibration-bias would have tended to produce the same left-right relationship for all *O*s, a result that was not apparent in the present data.

Assuming a generalized linear relationship between the two eyes, all the data for each *O* for each study were combined and the best-fitting solution was calculated by the method of least squares. Such a relationship is expressed as follows:

$$\begin{aligned}x_L &= m_{11}x_R + m_{12}y_R + m_{13} \\ y_L &= m_{21}x_R + m_{22}y_R + m_{23}\end{aligned}\tag{1}$$

which in matrix form becomes:

$$\begin{bmatrix} x_L \\ y_L \end{bmatrix} = M \begin{bmatrix} x_R \\ y_R \\ 1 \end{bmatrix}$$

where *M* is the 2 × 3 matrix:

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \end{bmatrix}$$

and: x_R, y_R = CIE trichromatic coördinates of the test color presented to the right eye.⁵ x_L, y_L = CIE chromatic coördinates of the mean matching color found for the left eye.

Luminance (CIE *Y*) was not included in this study because a comparative luminance-calibration for the two colorimeters used, unnecessary for the two studies previously reported, was not made. The more general relationship, which includes the luminance-dimension, is outlined in the Appendix.

When the *M* matrices had been solved by least squares, it began to look as though a simpler solution were possible, of the form:

$$\begin{aligned}x_L &= m_{11}x_R + m_{12} \\ y_L &= m_{21}y_R + m_{22}\end{aligned}\tag{2}$$

since the other coefficients seemed to approach zero. Equations relating the data were then calculated in this simpler form. Standard errors of prediction were computed for both forms of equations, and the average difference was only 0.0005 in CIE chromaticity units which is of sub-threshold visual magnitude. Consequently, the

⁴ Burnham, A colorimeter for research in color perception, this JOURNAL, 65, 1952, 603-608.

⁵ The CIE system of color specification is an international standard set forth by the Commission Internationale d'Éclairage. For a lucid discussion see D. B. Judd, *Color in Business, Science, and Industry*, 1952, 37-60; and Committee on Colorimetry O.S.A., *The Science of Color*, 1953, 254-316.

simpler equations are reported here. Equations were also computed in the reverse direction, predicting test-colors from the mean matching-colors. These are not reported here, but standard errors of prediction showed an average difference from those obtained with the first equations of 0.0008. The prediction-equations and standard errors for the two studies are as follows.

First study:

<i>O</i>		$\sigma_{L \cdot R}$	
<i>B</i>	$x_L = .9503x_R + .0133$ $y_L = .9281y_R + .0110$	.0141 .0183	[3a]
<i>RE</i>	$x_L = 1.0064x_R + .0058$ $y_L = 1.0279y_R - .0006$	.0112 .0118	[3b]
<i>N</i>	$x_L = 1.0114x_R - .0069$ $y_L = .9473y_R + .0125$	.0073 .0135	[3c]
<i>G</i>	$x_L = 1.2551x_R - .0659$ $y_L = 1.1776y_R - .0567$	.0168 .0157	[3d]
<i>D</i>	$x_L = 1.0163x_R - .0130$ $y_L = 1.0127y_R - .0151$	.0115 .0078	[3e]
<i>J</i>	$x_L = 1.0348x_R - .0114$ $y_L = 1.0845y_R - .0303$	.0247 .0192	[3f]

Second study:

<i>O</i>			
<i>B</i>	$x_L = .9412x_R + .0269$ $y_L = .8984y_R + .0383$	.0066 .0030	[3g]
<i>RE</i>	$x_L = 1.0974x_R - .0199$ $y_L = 1.0972y_R - .0242$	.0087 .0087	[3h]
<i>N</i>	$x_L = .9635x_R + .0142$ $y_L = 1.0142y_R + .0018$	.0044 .0073	[3i]
<i>C</i>	$x_L = .9934x_R - .0004$ $y_L = .9829y_R + .0001$	.0020 .0030	[3j]

Inspection of these equations revealed the possibility that a number of them approach the form:

$$\begin{aligned} x_L &= x_R \\ y_L &= y_R \end{aligned} \quad [4]$$

or zero difference between the two eyes. Tests of significance were made for departure of Equations [3] from the form of Equations [4] and these are summarized in Table I showing significance at the 5% and 1% levels for both directions of prediction.⁶ These tests show that significant differences were found in the color-sensitivity

⁶ A. M. Mood, *Introduction to the Theory of Statistics*, 1950, 291-297. The equation on p. 296 of this reference,

$$F = \frac{Q}{n\sigma^2/\sigma^2}$$

should be corrected to read:

$$F = \frac{Q}{2\sigma^2/\sigma^2}$$

for bivariate data.

of the two eyes of *O*s *D*, *RE*, and *G* in the first study and *B*, *C*, and *RE* in the second study. Where only one of the two equations showed significance at the 5% level, and neither at the 1% level, the difference was not considered significant. *B*, *N*, and *RE* were common to both studies. *N* and *RE* gave qualitatively consistent results in both studies, *N* being not significant in both and *RE* being significant. *B*, however, gave a different result in the two studies, not significant in the first but significant in the second.

Lack of significance would certainly be favored in the first study where there was ample but variable opportunity for local adaptation to the test- and matching-

TABLE I

SIGNIFICANCE OF THE DIFFERENCES OF THE *O*s' INTER-EYE SENSITIVITY
(Symbols in parentheses represent significance levels for reverse predictions.)

First study			Second study		
<i>B</i>	— (—)	* (*)	<i>B</i>	† (†)	† (†)
<i>RE</i>	* (*)	* (†)	<i>RE</i>	† (†)	* (*)
<i>N</i>	— (—)	— (—)	<i>N</i>	— (—)	* (*)
<i>G</i>	† (†)	† (†)	<i>C</i>	† (†)	† (†)
<i>D</i>	* (*)	† (†)			
<i>J</i>	— (*)	— (—)			

* Significant at the 5% level.

† Significant at the 1% level.

colors. If an *O* showed significant differences under the conditions of the first study, he could also be expected to show differences in the second study. If his inter-eye differences were large enough to appear even with the interference of local adaptation, they could almost certainly be expected to appear with a reduction of these effects. There were no exceptions to this expectation with the three *O*s who were common to the two studies.

Results. Plots of the actual match-specifications for particular test-colors show some interesting individual differences. These are shown in Figs. 1 and 2 in the CIE chromaticity plane, with vectors running from the specification of the test-color (right eye) to the match specification (left eye). *N* showed no significant inter-eye differences in either study, and the vectors representing his results show no obvious systematic trend. (This *O* reported that other researches on his eyes indicated significant inter-eye differences in luminance which could not be evaluated here.) *RE* showed significant differences in both studies, and plots of his data show generally that his left eye required more blue to complete a match. *B* showed no significant difference in the first study, but did show a significant inter-eye difference in the second study. These results are reflected in the vector plots; for the first study there is no consistent trend but for the second study there is a rather clear indication that his left eye required more yellow for a match. Vectors for all three *O*s tended to be longer for

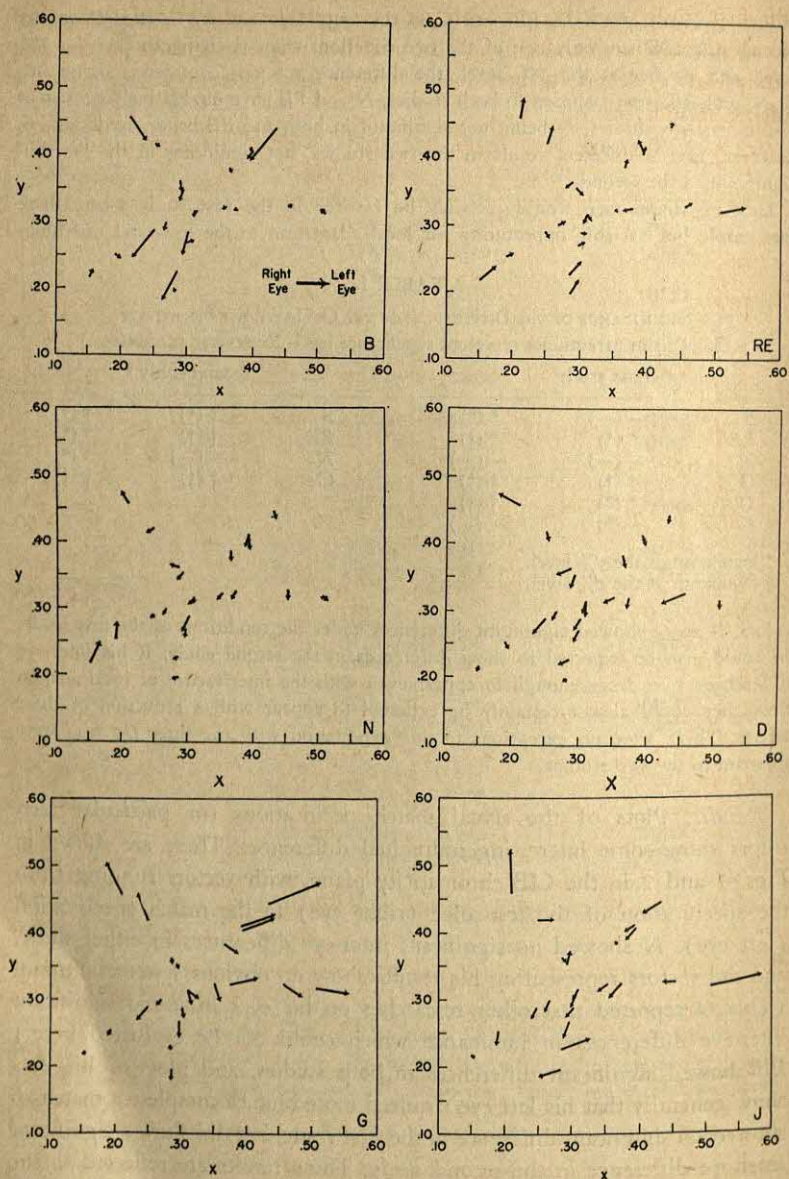


FIG. 1. CIE CHROMATICITIES OF COLOR MATCHES FOR *Os*' LEFT EYES COMPARED TO THE CHROMATICITIES OF THE TEST-COLORS PRESENTED TO THEIR RIGHT EYES
(First study: 6 *Os* each matched 22 colors.)

the first study, and this probably reflects interference due to the effects of local adaptation. Variability of the matches was also greater in the first study; hence, despite greater vector lengths, the differences were typically of less significance.

G, who observed only in the first study, showed a significant inter-eye

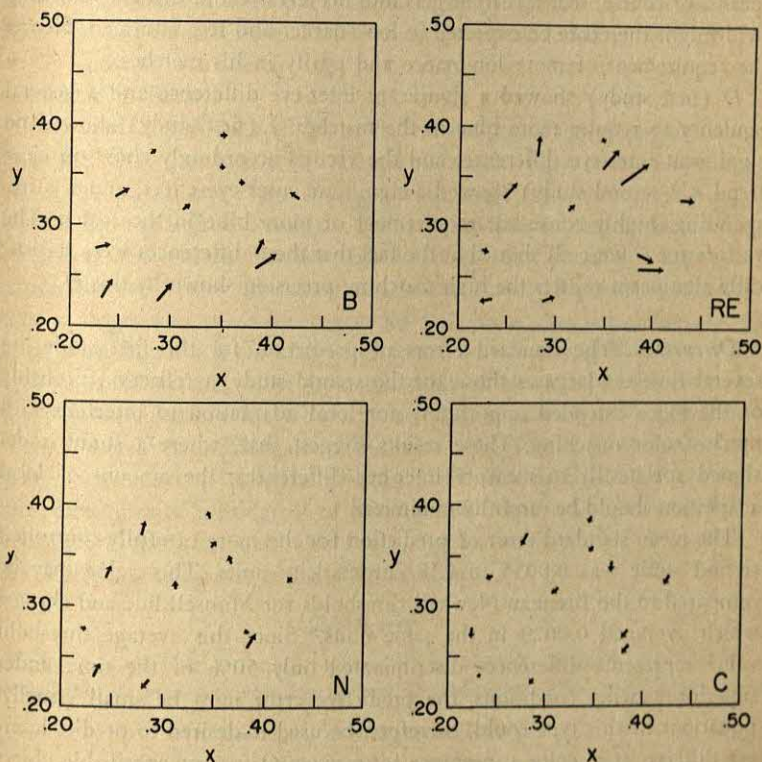


FIG. 2. CIE CHROMATICITIES OF COLOR MATCHES FOR Os' LEFT EYES COMPARED TO THE CHROMATICITIES OF THE TEST-COLORS PRESENTED TO THEIR RIGHT EYES (Second study: 4 Os each matched 12 colors.)

difference, but the trend of his results shows that his left eye required very consistently more purity to produce a match. In visual terms, there was an average increase of 2.1 Munsell chroma steps.⁷ It had also been noted during the study that he usually required more luminance for matches than

⁷ S. M. Newhall, Dorothy Nickerson, and D. B. Judd, Final report of the O.S.A. subcommittee on the spacing of the Munsell colors, *J. opt. Soc. Amer.*, 33, 1943, 385-418.

any of the other *O*s, and that left and right eye surrounds, equated for luminance, never appeared that way to him; he frequently remarked that the left surround was darker than the right surround. Without elaborate measurements, it was observed (and verified by him) that the pupillary area of his right eye is habitually about twice that of his left eye. This means, of course, that less light gets into his left eye and that the matching field might therefore be expected to look darker and less saturated. Hence, the requirement of more luminance and purity in his matches.

D (first study) showed a significant inter-eye difference and a general tendency to require more blue in the matches. *J* (first study) showed no significant inter-eye difference, and the vectors accordingly show no clear trend. *C* (second study) showed a significant inter-eye effect, and a corresponding, highly consistent requirement of more blue in the match. The vectors for *C* were all short, but the fact that these differences were statistically significant reflects the high matching precision shown by this *O*.

Discussion. The standard errors of prediction for the first study are several times as large as those for the second study, a reflection probably of the more extended opportunity for local adaptation to interfere with precise color matching. These results suggest that, where a study is designed specifically to measure inter-eye differences, the amount of local adaptation should be carefully minimized.

The mean standard error of prediction for the more carefully controlled second study was 0.0055 in CIE chromaticity units. This value may be compared to the Brennan-Newhall thresholds for Munsell hue and chroma which averaged 0.0029 in the same units.⁸ Since this average threshold value represents differences discriminated only 50% of the time under superior viewing conditions, the predictive error must be small visually. Equations of this type could, therefore, be used if desired to predict inter-eye differences in color appearance for a given *O* for any specifiable object color, once his equations have been determined. Such equations would, of course, be limited to a single individual since there can be great diversity among individuals in inter-eye differences. The size of the predictive error can also be compared to that reported in our second study where there were gross differences in adaptation of the two eyes. On the average that error was 0.0079 or about $\frac{1}{3}$ higher. Thus, the predictive error for inter-eye differences in color appearance is less when both eyes have the same

⁸ J. G. Brennan and S. M. Newhall, I.C.I. specifications of difference limens for Munsell hue, value, and chroma, *J. opt. Soc. Amer.*, 38, 1948, 696-702.

adaptation than when the two eyes are differently adapted. Such a result might be expected because of the relatively less complex viewing conditions.

Inter-eye differences in visual units can be estimated for all *O*s by comparing their differences on the CIE plots to those for *G* whose consistent requirement of more purity in the matches amounted on the average to 2.1 Munsell chroma steps. Estimates can be reliably made, though, only for *O*s showing statistically significant departures from zero difference as indicated in Table 1.

Conclusion. Significance of this numerical analysis of inter-eye differences lies in the technique rather than the specific results reported for a few *O*s, and in the possibility of predicting the inter-eye difference in color-appearance for any specifiable object, once equations have been derived from a relatively few colors. Further significance of this predictive approach accrues, however, from the future possibility of analyzing inter-eye differences in larger and more heterogeneous samples of the population. One of the most important clues to what colors are seen by individuals with major color defects has come from the study of isolated individuals with a major color defect in one eye but normal color vision in the other.⁹ Perhaps a careful sorting of individuals to find those who have large inter-eye differences will make it possible eventually to understand more fully the color-perceptions of all degrees of color defective eyes and the underlying mechanism as well. It could also make possible the definition of normal color-perception as a basic comparative criterion.

APPENDIX

Where it is possible to include luminance-measurements, as well as those for chromaticity, the prediction-equations assume the following more general form:

$$X_L = m_{11}X_R + m_{12}Y_R + m_{13}Z_R + m_{14}$$

$$Y_L = m_{21}X_R + m_{22}Y_R + m_{23}Z_R + m_{24}$$

$$Z_L = m_{31}X_R + m_{32}Y_R + m_{33}Z_R + m_{34}$$

which in matrix form becomes:

$$\begin{bmatrix} X_L \\ Y_L \\ Z_L \end{bmatrix} = M \begin{bmatrix} X_R \\ Y_R \\ Z_R \\ 1 \end{bmatrix}$$

⁹ Judd, Color perceptions of deuteranopic and protanopic observers *J. Res. Nat. Bureau of Standards*, 41, 1948, 247-271. In this article 37 cases of unilateral color defect are reviewed.

where M is the 4×3 matrix:

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix}$$

and: X_R, Y_R, Z_R = CIE tristimulus values of the test color presented to the right eye.
 X_L, Y_L, Z_L = CIE tristimulus values of the mean matching color found for the left eye. The M matrix can be solved by the method detailed in a previous study from this laboratory.¹⁰

¹⁰ W. L. Brewer, Fundamental response functions and binocular color matching, *J. opt. Soc. Amer.*, 44, 1954, 207-212.

THE INFLUENCE OF FIGURAL INSPECTION ON THE AUTOKINETIC ILLUSION

By J. E. CONKLIN, Minneapolis-Honeywell Aeronautical Division

Several recent studies suggest a relation between figural inspection and the perception of apparent movement. Research has demonstrated: (a) that continuous fixation of a figure or of apparent movement distorts optimal stroboscopic movement; (b) that prolonged stimulation of apparent movement produces modifications in the subsequent perception of familiar figures; and (c) that sustained figural inspection reduces the extent of autokinetic movement.¹ The experiments reported here on the influence of figural inspection on the autokinetic illusion are designed to test several specific deductions derived from Köhler and Wallach's cortical satiation theory.²

It is probable that autokinetic movement is not the result of a single process, but rather, a phenomenon determined by several variables. The apparent movement, for example, can be initiated by eye-movements.³ Retinal and cortical processes are also involved in the illusion, *e.g.* streaming, satiation.⁴ According to the writer's interpretation of satiation theory, the autokinetic illusion is determined by the flow of figural currents in the visual projection area. Figural currents, instantaneously set up by the spot of light, decrease the conductance of the cortical media by polarization and

* Received for publication October 1, 1956. This study, financed by a grant from the National Science Foundation, was conducted at Montana State College. The author's thanks are due Marlys Jolley, Juanita Nickerson, Flora Wright, and Barbara Hill for their assistance on this study.

¹ B. H. Deatherage and M. E. Bitterman, The effect of satiation on stroboscopic movement, this JOURNAL, 65, 1952, 108-109; May W. Brenner, Continuous stimulation and apparent movement, *ibid.*, 66, 1953, 494-495; Deatherage, Figural after-effects in stroboscopic movement, *ibid.*, 67, 1954, 259-267; M. B. Shapiro, A preliminary investigation of the effects of continuous stimulation on the perception of 'apparent motion,' *Brit. J. Psychol.*, 45, 1954, 58-67; R. H. Christman, Figural after-effects utilizing apparent movement as an inspection figure, this JOURNAL, 66, 1953, 66-72; N. H. Livson, After-effects of prolonged inspection of apparent movement, *ibid.*, 66, 1953, 365-376; Ward Edwards and R. S. Crutchfield, Differential reduction of autokinetic movement by a fixated figure, *J. exp. Psychol.*, 42, 1951, 25-31; Crutchfield and Edwards, The effect of a fixated figure on autokinetic movement, *ibid.*, 39, 1949, 561-568.

² Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. phil. Soc.*, 68, 1944, 269-357.

³ Alex Skolnick, The role of eye movements in the autokinetic phenomenon, *J. exp. Psychol.*, 26, 1940, 373-393.

⁴ J. P. Guilford and K. M. Dallenbach, A study of autokinetic sensation, this JOURNAL, 40, 1928, 83-91; Guilford, Autokinesis and the streaming phenomena, *ibid.*, 40, 1928, 410-417.

electrotonus. After a period of time (autokinetic latency) the figural current follows a new path of lesser resistance—changing the phenomenal position of the autokinetic stimulus. To account completely for autokinetic movement, however, additional assumptions are necessary: (a) the cortical area is in a semipermanent and irregular state of satiation; and (b) the eye cannot maintain constant fixation due to physiological nystagmus thereby releasing a nonsymmetrical, statistical distribution of impulses at the visual cortex.⁵ Without these assumptions satiation theory would predict that with continued fixation, the autokinetic stimulus would increase in size and become fainter.

If this explanation is correct, experimental conditions systematically inducing cortical satiation should alter the direction, rate and extent of autokinetic movement. Furthermore, since satiation is primarily a function of contour gradients (as Köhler and Wallach's theory suggests), satiation figures with a greater perimeter to area ratio should enhance the after-effect. This argument points to several deductions: (1) Autokinetic movement will travel away from a satiated area. (2) The rate and extent of apparent movement will vary directly as the angular size of the satiated area. (3) Inspection circles will have a greater effect on the illusion than squares subtending the same visual angle. (4) There will be no significant differences in the after-effect for filled (blackened-in) or outlined inspection-figures.⁶

METHOD AND PROCEDURE

Apparatus. Quantitative measures of latency, direction, extent, and rate of autokinetic movement are obtained by recording *O*'s open-loop, pursuit tracking response.⁷ The procedure is to train *O* to follow a moving spot of light with a two-dimensional control under conditions where apparent movement is absent and then present him with the autokinetic situation. The tracking task is similar and yields relatively precise measures of the several autokinetic criteria.

A spot of green light ($\frac{1}{2}$ in. in diameter) is projected from an oscilloscope to the exact center of a 40×40 -in. screen located 100 in. away.⁸ For training purposes, the rate at which the spot is moved is controlled by condenser-resistor charging circuits and the direction of movement can be regulated separately in the vertical

⁵ C. E. Osgood and A. W. Heyer, Jr., A new interpretation of figural after-effects, *Psychol. Rev.*, 59, 1952, 98-118.

⁶ For a review of the literature and a statement of the problem see J. E. Conklin, The influence of figural inspection on visible movement: I. The figural after-effect, *Northwest Science*, 30, 1956, 77-98; II. The autokinetic illusion and visual rate perception, *ibid.*, 30, 1956, 133-145.

⁷ Conklin, A note on the measurement of autokinetic movement, this JOURNAL, 68, 1955, 144-145; Quantitative determination of autokinetic movement: Rate, direction, extent, and latency, *Northwest Science*, 30, 1956, 61-69.

⁸ J. E. Conklin, A. Baldwin, and R. H. Brown, Apparatus for measuring the threshold for the visual discrimination of direction of movement, this JOURNAL, 66, 1953, 289-294.

and horizontal dimensions. The duration of the training and autokinetic stimulus is determined by cam-microswitch circuits furnishing intervals of 0.5, 1, 2, 4, 8, and 16 sec.

Tracking is accomplished with a war surplus two-dimensional gun-fire control. The vertical and horizontal Selsyn generators have been replaced by high precision linear potentiometers and wired to an oscillograph and battery. Control movements in azimuth and elevation are recorded on a constant speed, two-channel Sanborn recorder. The control has no influence on the movement of the spot but can be pointed to any position on the projection screen. The sensitivity of the recorder was so adjusted that the stylus of the recorder is displaced 20 mm. when the control is rotated 18.4° in azimuth or elevation. This represents the angle subtended when the control is moved from the center position and pointed to the edge of the projection screen, *i.e.* 20 in.

In one corner of the large experimental room, a dark booth was constructed $65 \times 140 \times 93$ in. in width, length, and height respectively. A bite-board, firmly anchored to the concrete wall, extends at a fixed height locating *O*'s eyes 72 in. from the center of the projection screen. An adjustable chair is used to take into account individual differences in trunk length. The projection screen is mounted on $\frac{1}{2}$ -in. plywood fastened to another concrete wall in the darkroom. Hinged to this is a second plywood square constructed to hold satiation figures drawn on 40×40 -in. white cardboard. The control panel is located behind the darkroom which is illuminated by a 10-w. lamp. During preliminary training, the curtains are parted on one side of the darkroom and the overhead lights turned on so that the visual field will be sufficiently structured to prevent the occurrence of autokinetic movement.

Procedure: (a) Preliminary practice. *O* is instructed that the experiment consists of learning to follow a moving spot of light in the dark by the 'feel of the control.' He is informed that practice sessions will be given in following vertical as well as horizontal movements under conditions of semi-darkness before the main experiment. Instructions are also given for the use of the bite-board and on the operation of the control.

Before the first practice trial, *O* is instructed to move the control (without looking at it) from the center position to the right until he thinks it points to the edge of the screen. *E* gives immediate knowledge of results to correct over- or undershooting of the edge. *O* is then asked to return the control to center, and then to a position that is halfway between the center and edge of the screen. During these procedures, *E* constantly reminds *O* not to look at the control while responding but rather to learn positioning by the 'feel of the control.' These procedures are similar for control movements to the left, and up and down.

Practice sessions consist of training *O* to follow rates of movement ranging from 0.5° to 4.0° of visual angle per second. The time-exposures employed are 4, 8, and 16 sec. For each time-interval, the spot moves from the center to the edge of the screen, or to a point halfway between the center and edge. Target rates, therefore, are 2° and 4° , 1° and 2° and 0.5° and 1° of visual angle per second, respectively. Practice is continued for movements of the spot to the right, left, up and down until *O* is tracking adequately as indicated by his response on the recorder. The learning of this task is relatively simple.

After the practice session, *O* is given a 5-min. rest in dim illumination (illumination from a 15-w. lamp mounted above *O*'s head and spotted on the screen automatically between trials). During this period, the overhead lights of the experimental room are extinguished, and the curtains are fastened to prevent extraneous light from entering the darkroom. Afterward, *O* is instructed to fixate the small cross on the stimulus-card without moving his eyes (see Fig. 1). After fixating 1 min., *O* is informed that his tracking ability for two-dimensional movement will now be tested in complete darkness. *E* removes the stimulus-card by opening the hinged panel covering the projection screen, and then presents the autokinetic stimulus for a 16-sec. exposure. Four trials made in rapid succession follow each fixation-period. After each trial, *O* is instructed to return his control to the center position—minor corrections are verbally initiated by *E*. The next autokinetic trial begins immediately following a 'ready' signal. Control movements in train and elevation are recorded for each autokinetic trial. From these recordings latency, direction, maximal displacement, and average rate of autokinetic movement are obtained.

Satiational stimulus-figures. The independent I-figure variables studied were size, filled, and outlined I-figures, and the orientation of the I-figures. Twelve stimulus-figures were drawn on 40 × 40-in. white cardboard. Circles and squares subtending visual angles of 3°, 6°, and 9° were employed. Half of the stimulus-figures were drawn in outlined form and half were completely blacked in. An additional variable was the orientation of the stimulus-card during fixation, e.g. Fig. 1 shows an orientation to the left. Stimulus-card orientations were varied to the left, right, up, and down.

Observers. Six undergraduates (men) of Montana State College served as *O*s in the first experiment. They were paid for their services. In the second experiment, 96 volunteer *O*s were employed; 27 women and 69 men, all students in general psychology. The *O*s were neither informed as to the purpose of the experiment, nor instructed on the nature of the autokinetic illusion.

EXPERIMENT I

In the first experiment, 6 *O*s served under all 48 experimental combinations. They reported to the laboratory for 13 consecutive days (excluding weekends) for sessions lasting from 45 to 60 min. The first day was devoted entirely to training. Thereafter, throughout the course of the experiment, the first 15 min. of each session was devoted to practice. On any given experimental day, *O* was subject to four 1-min. figural inspection-periods followed by four autokinetic trials of 16 sec. The orientation of the stimulus-cards was constant for any given day but determined by a table of random numbers for each *O*. The remaining experimental variables were also randomized separately for each *O*. Restrictions in the randomization consisted of insuring that all experimental variables were presented for all orientations of the stimulus-card.

Results. Vertical and horizontal components of the autokinetic illusion were analyzed separately. Three autokinetic criteria (latency, rate, and displacement) were subject to an analysis of variance, whereas the direction of autokinetic movement was individually assayed for each orientation of the I-figure. Double interactions were tested with an error-term com-

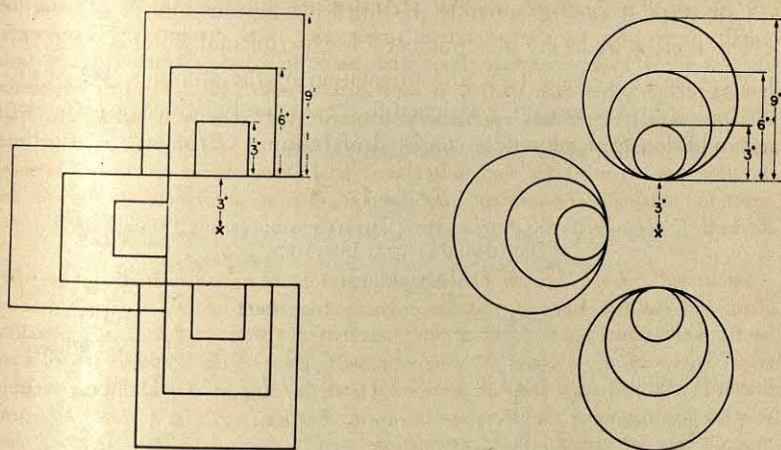


FIG. 1. INSPECTION-FIGURES ORIENTED TO THE LEFT

prised of the residual and higher order interactions. The double interactions of *O*s by main-effect were employed to test the I-figure variables, viz. size, shape, filled-unfilled, orientation, and trials.

I-figure variables are found to have no significant bearing on the direction, latency, rate, or displacement of autokinetic movement, *i.e.* none of the main effects is significant at the 1% level of confidence. In addition, there are no significant interactions among I-figure variables. Differences observed between vertical and horizontal components of the autokinetic movement are, however, highly significant at the 1/10% level of confidence. Individual differences among *O*s are reflected in all the quantitative criteria ($p < 0.001$).

Table I shows a decrease in the rate and displacement of autokinetic movement as a function of trials following figural inspection. These results are significant for both vertical and horizontal components of the illusion ($p < 0.001$).

The results of this study did not support the predictions that were made on the basis of satiation theory. The hypothesis, for example, that I-figure variables were systematically related to autokinetic movement was not

demonstrated. That this was due to sequential effects in the experimental design, *i.e.* using *O*s to serve under all experimental treatments, was suspected. A second experiment, therefore, was conducted to investigate this possibility.

EXPERIMENT II

In the second experiment only the filled-in stimulus-figures were employed, thereby reducing the number of experimental combinations to 24, *viz.* size (3), shape (2), and orientation of the stimulus-cards (4). Ninety-six *O*s were randomly assigned to the 24 combinations, 4 *O*s under each condition. The procedure consisted of training *O* for 30 min. in the

TABLE I
MEANS ON THREE AUTOKINETIC CRITERIA FOR SUCCESSIVE TRIALS
FOLLOWING FIGURAL INSPECTION
(Experiment I)

Criteria		Movement				Mean
		1	2	3	4	
Horizontal	Latency (sec.)	6.9	7.3	7.7	8.1	7.50
	Rate (min./sec.)	12.2	10.0	9.9	8.6	10.15
	Displacement (min.)	115.7	95.7	91.0	78.2	95.17
Vertical	Latency (sec.)	2.4	1.8	1.8	1.7	1.89
	Rate (min./sec.)	29.9	25.9	24.5	24.8	27.28
	Displacement (min.)	364.7	316.5	307.8	306.7	323.93

manner described previously and then presenting him with six, 1-min. fixation-periods, following each with four, 16-sec. autokinetic trials. Three inspection periods were of the stimulus-card and three where *O* fixated a spot of light identical in all respects to the autokinetic stimulus. Spot and figure inspections were randomly sequenced for each *O*.

Results. Data derived from this investigation were analyzed as follows: (a) the differential effect of spot and figural inspection on autokinetic movement was determined for all criteria; (b) the significance of the direction of autokinetic movement was computed for each I-figure orientation; (c) an analysis of variance was performed for the criteria of latency, rate and displacement following figural inspection; and (d) vertical and horizontal components of the illusion were analyzed separately.

The results of this study parallel the negative findings of Experiment I. The only significant results were the differences observed between vertical

and horizontal components of the autokinetic illusion (see Table II). Autokinetic movement following figural inspection was not significantly different than the movement observed when merely the spot was fixated. In addition, there was no decrement in the rate and displacement of autokinetic movement as a function of trials following the fixation-period as observed in Experiment I. It is likely that the trials-effect was an extraneous result of the experimental design, related to factors of boredom and fatigue.

DISCUSSION

The results of these two studies are characterized by negative findings. There is no evidence to indicate that prolonged fixation of a figure pro-

TABLE II
MEANS ON THREE AUTOKINETIC CRITERIA FOR HORIZONTAL AND VERTICAL
MOTION FOLLOWING FIGURE- AND SPOT-FIXATION
(Experiment II)

Criteria	Fixation	Horizontal	Vertical
Latency (sec.)	figure	3.9	3.5
	spot	3.8	3.6
Rate (min./sec.)	figure	9.8	19.8
	spot	9.7	22.1
Displacement (min.)	figure	94.5	224.3
	spot	100.7	248.9

duces a subsequent modification of the autokinetic illusion. According to satiation theory, systematic variation of satiation figures should have a definite influence on the direction of autokinetic movement; if not, the rate and extent of the movement should be reduced when the apparent movement is in the direction of the satiated area. Table III summarizes these data for horizontal and vertical components of autokinetic movement with each figure orientation employed. The differences between left-right and up-down measures are neither significant, nor are they consistently in the predicted direction.

The negative findings of these two studies seriously question the validity of the satiation theory as an explanation for the autokinetic illusion. The phenomenon may be primarily determined by an interaction of peripheral factors, such as eye-movement, adaptation, and streaming. This interpretation is not to be construed as supporting an untenable position that the autokinetic illusion is determined solely by retinal events, but rather, that the apparent movement is probably initiated by peripheral factors. Why the apparent movement is influenced in the presence of social stimulation, for example, is yet to be explained.

SUMMARY

Two experiments were conducted to determine the effect of figural inspection on the subsequent perception of the autokinetic illusion. I-figures varying in size, shape, filled and unfilled, and orientation in the visual field were fixated by Os for 1 min. prior to the autokinetic situation. Quantitative

TABLE III
RIGHT-LEFT AND UP-DOWN MEANS FOR HORIZONTAL AND
VERTICAL FIGURE-ORIENTATIONS

	Criteria	Left		Right		Up		Down	
		R	L	R	L	U	D	U	D
Experiment I	Latency (sec.)	4.9	5.8	4.6	5.7	1.3	2.7	1.6	3.3
	Rate (min./sec.)	17.3	13.9	20.6	17.1	27.4	30.2	27.9	27.0
	Displacement (min.)	127.3	98.1	120.8	125.9	337.1	340.8	339.2	249.5
Experiment II	Latency (sec.)	5.3	5.4	5.3	6.3	3.8	4.3	4.8	6.7
	Rate (min./sec.)	19.3	17.2	18.1	20.9	25.3	28.5	21.6	20.8
	Displacement (min.)	118.9	120.8	120.6	127.1	268.7	231.5	218.8	149.4

measures of direction, latency, rate and displacement of autokinetic movement were obtained by an open-loop, pursuit tracking device. All Os received practice in tracking a moving spot of light with the apparatus before the main experiments.

The results were negative. It was found that there was no relation between I-figure variables and autokinetic movement. It was concluded that retinal events play a predominant role in the apparent movement since cortical satiation appears to be unrelated to the phenomenon.

SPATIAL AND TEXTURAL CHARACTERISTICS OF THE GANZFELD

By WALTER COHEN, University of Buffalo

There have been several partially successful attempts to study the effects of uniform stimulation over the entire visual field. Metzger has reported that under low illumination the entire visual field (*Ganzfeld*) appears as "a mist of light which becomes more condensed at an indefinite distance."¹ When intensity is increased, the fog becomes a filmy surface separated from *O* by empty space. The distance of the surface is judged to be about the same as its objective distance. With further increase in level of illumination, the distance is judged to be greater, and the filmy curved surface is transformed into a flat plane surface with microstructure. It is difficult to interpret these results since uniformity of stimulation was not achieved by Metzger. Under high illumination, the physical texture of the surface was perceived.

Gibson and Waddell, using a translucent globe, found that under high illumination their *O*s described a "sea of light" with something vaguely surface-like in front of the face.² Their naïve *O*s did not consistently report the impression of space-filling fog. Gibson and Waddell attribute the occasional reports of other qualities to central processes rather than to peripheral stimulation.

There have been two basic sources of inhomogeneity in all previous studies dealing with the *Ganzfeld*. First, physical microstructure was present. In addition, parts of *O*'s face prevented uniform stimulation of the entire retina—for example, the nose must have produced a shadow on the peripheral nasal area of the visual field. The apparatus developed for the present investigation provided an adequate uniform *Ganzfeld* by eliminating both these sources of inhomogeneity. Although additional information was collected, this paper will be limited to the 'mode of appearance' and spatial characteristics of the *Ganzfeld*. The following three questions guided the design of the experiment: (1) In what manner is the appearance of the homogeneous *Ganzfeld* dependent upon intensity and wavelength? (2) How does the appearance of the inhomogeneous

* Received for publication October 29, 1956. The writer is indebted to Professor Carl R. Brown for his assistance. The study is based on a portion of a doctoral dissertation submitted to the University of Michigan.

¹ Wolfgang Metzger, Optische Untersuchungen am Ganzfeld. II. Zur Phänomenologie des homogenen Ganzfelds, *Psychol. Forsch.*, 13, 1930, 6-29.

² J. J. Gibson and Dickens Waddell, Homogeneous retinal stimulation and visual perception, this JOURNAL, 65, 1952, 263-270.

Ganzfeld differ from that of the homogeneous *Ganzfeld*? (3) How are these differences in appearance related to stimulus-distributions?

METHOD

Apparatus. A new method was devised which would produce a uniform *Ganzfeld* as well as permit the introduction of a spot into the field. The principle utilized was similar to that of the photometric sphere—as a result of direct illumination of a

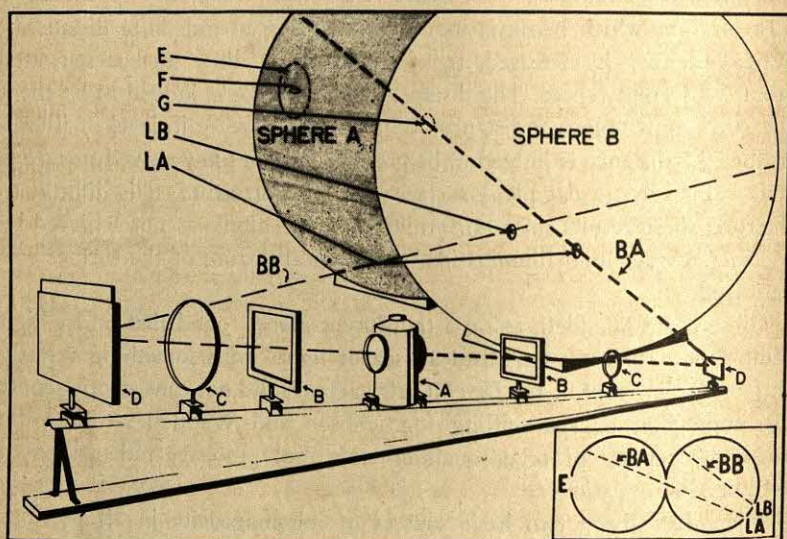


FIG. 1. DIAGRAM OF APPARATUS

A, light-source; B, filter; C, lens; D, mirror; E, mask; F, eye-piece; G, 8-cm. aperture; BA, beam for sphere A; BB, beam for sphere B; LA, opening for BA; LB, opening for BB.

small part of the inner surface, there is indirect uniform illumination of the entire surface of the sphere.

Two spheres, each having a diameter of 1 m., were joined together (Fig. 1). The smooth inner walls were sprayed with Benjamin Moore Sphere Paint to produce a diffuse reflecting surface of high albedo and no visible texture at viewing distance. The O looked into Sphere A, and the light entered by way of two openings in Sphere B. At the junction of the two spheres was a circular aperture 8 cm. in diameter. A special mask was made to fit the contours of O's face. The mask formed a removable section of the wall of Sphere A and was used as a head-rest. Each mask had an opening which permitted unobstructed monocular vision. When O looked to the right, no part of his mask or face was visible. The left eye was kept in complete darkness. Since the opening of the mask was directly opposite the center of the 8-cm. aperture, O's line of sight was perpendicular to the plane of the aperture.

A beam of light was projected onto the inner surface of Sphere A, 20 cm. above

O's eye. Sphere *B* was uniformly illuminated by a beam of light that could be independently varied. By use of appropriate filters, the luminosity of both spheres was so controlled as to produce either equality or some desired difference. When the luminous flux-density in the two spheres was made equal, the aperture ceased to exist optically, and a uniform *Ganzfeld* was produced. When the luminous flux in the two spheres differed as to dominant wavelength, purity, or intensity, inhomogeneity was produced. Light coming from the aperture-area was different from that coming from the rest of the field, and *O* was confronted with a simple inhomogeneity in an otherwise uniform field.

The two beams of light originated from the same source. With appropriate lenses and mirrors, the beams were directed through two 5-cm. openings in Sphere *B*. These openings were not visible to *O*s looking into Sphere *A*. Neutral and colored filters altered the intensity and composition of the light. Chromatic illumination was obtained by the use of the following Wratten filters: (1) red (608 m μ) CC-50R, excitation-purity 21%; (2) green (554 m μ) CC-40G, excitation-purity 21%; (3) green (554 m μ) CC-30G, excitation-purity 16%; (4) blue (461 m μ) CC-30B, excitation-purity 20%. The following notation will be used to designate the manner of illumination: The chromatic filters will be designated by the letters *R*, *G*, *GX*, and *B* respectively. The letter *A* will signify the absence of a chromatic filter. The luminance will be expressed in millilamberts. Thus, (*R*2.1-*A*4.9) signifies a field having a dominant wavelength of 608 m μ and a brightness of 2.1 m L ., with an achromatic spot having a brightness of 4.9 m L .

Observers. Thirteen men and three women served as *O*s. The median age of the group, which was made up of three undergraduates, ten graduate students, and three faculty members of the University of Buffalo, was 25 yr. Although all specialized in psychology, they were naïve as to the purpose of the present study. All had normal color vision, but only six had normal visual acuity.

Procedure. Each *O* was studied in an average of five 1-hr. sessions. A complete record of *O*'s responses was made with a tape-recorder. Tasks of two kinds were performed: (1) *O* was required to compare two situations with respect to fog-density, distance, distinctiveness of figure, hue, and saturation. Each of the situations was presented for 20 sec., with an interval of 5 sec. between the two members of each pair and an interval of 15 sec. between pairs. (2) *O* remained in the experimental situation for 3 min. to report on the effects of adaptation. Each situation was followed by a 2-min. rest-period.

The order of presentation was counterbalanced to control for time-error, practice, and adaptation. Each *O* was presented with approximately half of the experimental situations, thus providing an *N* of 8 for each. There were five basic kinds of visual field with which the *O*s were presented: (1) homogeneity; (2) inhomogeneity due to a difference in chromaticity between field and spot; (3) inhomogeneity due to a difference in intensity; (4) inhomogeneity due to a difference both in intensity and chromaticity; (5) inhomogeneity due to a difference in purity.

RESULTS AND DISCUSSION

(1) *Homogeneity.* The most characteristic description of the field for all

Os was "fog-like." Reports as to density and degree of immersion varied. The following phrases are representative: "A diffuse fog." "A hazy insipid yellow." "A gaseous effect." "A milky substance." "Misty, like being in a lemon pie." "Smoky."

The *Os* of Gibson and Waddell reported other impressions with considerable frequency, a disparity which may be traced to procedure as well as to apparatus.³ Gibson and Waddell used a questionnaire and a training procedure which may have influenced their *Os* to seek something other than fog. Failure to control for facial shadows may account for the reports of a boundary. During the present investigation, there were a few descriptions of a "cracked ice effect" or a "web-like structure" which could be eliminated by slight adjustments in the illumination of Sphere *B*. It may be that gradients of intensity which are insufficient for the formation of a spot result in other forms of phenomenal inhomogeneity.

The homogeneous *Ganzfeld* was reported to be "close at hand" by all *Os*. Only one consistently described the field as being more than 6 in. away, and the modal judgment of distance was 2 in. The fog itself was, however, seen to extend for an indefinite distance. Several *Os* described an experience of complete immersion in the fog, and, in a few instances, space was internalized, described as being "inside the head." Werner has described a similar internalization of sound.⁴ There was a high degree of uncertainty as to distance and most *Os* were extremely hesitant about making distance-judgments.

A nonparametric technique gave no indication, within the range of 1.8-4.9 mL., that either fog-density or fog-distance was significantly related to intensity of illumination ($P > 0.10$),⁵ a result which does not correspond to Metzger's findings.⁶ This discrepancy may be accounted for by the absence of surface-texture in the present experiment. The 'blue' field (*B2.4-B2.4*) was consistently judged ($P < 0.05$) as less dense than either the 'red' field (*R2.1-R2.1*) or the achromatic field (*A2.4-A2.4*). In addition, the *Os* reported feeling less immersed in the 'blue' field, which suggests that receptors sensitive to long wavelength are more densely concentrated than those sensitive to short.

During the course of adaptation, 5 of the 16 *Os* reported a complete cessation of visual experience. Two reported "black-out" almost every time

³ Gibson and Waddell, 268-269.

⁴ Heinz Werner, Untersuchungen über Empfindung und Empfinden, *Z. Psychol.*, 114, 1929, 152-166.

⁵ W. J. Dixon and A. M. Mood, The statistical sign test, *J. Amer. statist. Ass.*, 41, 1946, 557-566.

⁶ Metzger, 14.

they were confronted with prolonged, homogeneous stimulation. This was a unique experience which involved a complete disappearance of the sense of vision for short periods of time, and not simply the presence of a dark, undifferentiated visual field. The following description is representative: "Foggy whiteness, everything blacks out, returns, goes. I feel blind. I'm not even seeing blackness. This differs from the black when lights are out." It may be conjectured that the perceptual mechanism has evolved to cope with a differentiated field, and, in the absence of differentiation, there is a temporary breakdown of the mechanism (which might be detected electrophysiologically). Anxiety and fear of blindness under somewhat similar conditions also were reported by Hochberg, Triebel, and Seaman.⁷

Although uniform proximal stimulation does not provide a basis for articulation, the separation of the field into a fog in front of a ground occurs. This effect may be related to the process of figure-ground segregation. It should be noted that when hue is experienced in the chromatic *Ganzfeld*, it is seen as a phenomenal characteristic of the fog, whereas the ground behind the fog may appear as neutral.

(2) *Inhomogeneity*. Introduction of inhomogeneity reduced fog-density and increased fog-distance ($P < 0.05$) in most instances. When figure-ground segregation was poor because of the absence of steep gradients of intensity, the field, as fog, continued to be seen as close, while the figure was seen as far away. Thus, the 'normal' relationship between figure and ground was reversed. Distance-judgments for the figure under these conditions ranged from 5 in. to 6 ft., with a mode of 2 ft. Increasing the difference in intensity between field and spot resulted in a closer figure and a more distant ground. When the difference in intensity was maximal, judgments of figural distance ranged from "next to the eye" to 3 ft. away, with a median of 6 in. The ground usually was judged to be a few inches behind the spot, although a few *O*s judged it to be several feet behind the spot. Each *O* was consistent in his judgments of distance. There were, however, marked differences among *O*s in the range of values used to describe the distance of the figure. The terms 'close' and 'far' varied considerably in meaning for different *O*s. For one, an apparent distance of 5 in. was considered far, while for another an apparent distance of 3 ft. was considered close.

Inhomogeneity due either to intensitive or to chromatic differences between field and spot reduced fog-density and increased fog-distance in most

⁷ J. E. Hochberg, William Triebel, and Gideon Seaman, Color adaptation under conditions of homogeneous visual stimulation (*Ganzfeld*), *J. exp. Psychol.*, 41, 1951, 153-159.

cases. There were, however, no consistent differences ($P > 0.10$) between homogeneous and inhomogeneous fields when (1) differences in intensity between field and spot were at about 1 *jnd*, or (2) inhomogeneity was due to a difference in purity between field and spot. In both cases, the spot was unstable and often disappeared for brief intervals.

The results of comparisons between inhomogeneous fields are shown in Table I. Increasing the difference in intensity between field and spot tended to decrease the apparent density and to increase the apparent distance of the

TABLE I
RELATIVE FOG-DENSITY AND FOG-DISTANCE OF THE GROUND-
AREAS OF INHOMOGENEOUS FIELDS

Situations compared		Density-judgment				Distance-judgment				
	1	2	1-greater	Equal	2-greater	P*	1-nearer	Equal	2-nearer	P*
a	(A2.1-A2.0)	(A2.1-A1.8)	6	2	0	<0.03	5	3	0	<0.05
b	(A2.1-A1.8)	(A2.1-A0.0)	8	0	0	<0.01	6	2	0	<0.03
c	(A2.1-A2.2)	(A2.1-A2.4)	7	1	0	<0.03	5	3	0	<0.05
d	(A2.1-A2.4)	(A2.1-A4.9)	8	0	0	<0.01	7	1	0	<0.03
e	(A2.1-A2.4)	(A2.1-A1.8)	2	5	1	>0.25	1	6	1	>0.25
f	(R2.1-A2.2)	(R2.1-R2.2)	2	5	1	>0.25	1	6	1	>0.25
g	(R2.1-A2.0)	(R2.1-R2.0)	1	6	1	>0.25	2	6	0	>0.25
h	(R2.1-A1.8)	(R2.1-R1.8)	3	4	1	>0.25	4	2	2	>0.25
i	(G2.7-A2.4)	(G2.7-G2.4)	3	3	2	>0.25	3	4	1	>0.25
j	(A2.1-R2.1)	(A2.1-G2.1)	0	3	5	<0.05	2	3	3	>0.25
k	(A2.1-R2.1)	(A2.1-B2.1)	0	2	6	<0.03	0	3	5	<0.05
l	(A2.1-G2.1)	(A2.1-B2.1)	1	4	3	>0.25	1	5	2	>0.25

* Sign test.

fog (*a, b, c, d*). The direction of the difference was not relevant (*e*). There was no consistent difference between those situations involving both chromatic and intensive inhomogeneity and those involving only intensive inhomogeneity (*f, g, h, i*). The appearance of the fog of an achromatic field was dependent upon the dominant wavelength of the spot (*j, k, l*). The fog was reported as denser and closer when the dominant wavelength of the spot was short than when it was long.

An amorphous field, as well as an amorphous figure, appears to characterize those situations in which gradients of intensity are absent. The assertion that gradients of intensity are more effective than chromatic gradients in reducing the amorphous characteristics of the field would not, however, be completely warranted. Such a conclusion would require the comparison of fields in which the two types of gradient were equated as to magnitude. Since such an equation would be arbitrary, and articulation of the field is dependent upon the magnitude of the stimulus-gradients, no direct test of this hypothesis is possible.⁸

The mode of appearance of the field seems to be related to the definite-

⁸ Kurt Koffka and M. R. Harrower, Colour and organization, *Psychol. Forsch.*, 15, 1931, 145-275.

ness of figure-ground segregation. With increased sharpness of boundary, there is greater definiteness of surface. Gibson and Dibble, in their study of visual surface, found that the steepness of gradients of luminous intensity was a determinant of surface-hardness.⁹ The same was found to be true in the present study to a limited extent. Even when steep gradients of intensity were present, fog was reported more than 50% of the time. In only a few cases did the Os actually report the ground as appearing hard, but the figure was frequently described as hard when differences in intensity between field and spot were large.

An expansion of the field was experienced when steep gradients were introduced. Koffka suggested that contraction and expansion of space may be a function of 'forces' related to articulation of the field.¹⁰ In accordance with his formulation, well-articulated fields are perceived as being at a greater distance than those which are poorly articulated.

In most cases, the figure was judged to be about one and one half times as large as its 'real' size. It has been suggested that there is an invariant relationship between phenomenal size and phenomenal distance.¹¹ If there were such a relationship, the underestimation of the apparent distance of the figure should have been accompanied by an underestimation, rather than an overestimation, of its apparent size. When steepness of the intensive gradients was increased, the figure appeared closer, but there was no systematic change in apparent size. The apparent distance changed without a change either in retinal size or in apparent size. These results are consistent with the conclusion of Kilpatrick and Ittleson that the invariant relationship between size and distance is restricted to special conditions and does not apply to those situations in which the field is poorly articulated.¹²

SUMMARY

The present investigation deals with visual perception when the entire visual field is uniform and when a small differentiated area is introduced into such a field.

The most representative description of the homogeneous *Ganzfeld* is that of close, impenetrable fog. The experience is a unique one, and most Os have difficulty in describing the field in terms usually associated with visual phenomena. In some instances, temporary cessation of visual ex-

⁹ J. J. Gibson and F. N. Dibble, Exploratory experiments on the stimulus conditions for the perception of a visual surface, *J. exp. Psychol.*, 43, 1932, 414-419.

¹⁰ Koffka, *Principles of Gestalt Psychology*, 1935, 119.

¹¹ E. G. Boring, Visual perception as invariance, *Psychol. Rev.*, 59, 1952, 141-148.

¹² F. P. Kilpatrick, and W. H. Ittleson, The size-distance invariance hypothesis, *Psychol. Rev.*, 60, 1953, 223-231.

perience is reported. Variations in wavelength rather than intensity of illumination alter the density and distance of the fog.

A differentiated area in the *Ganzfeld* changes its phenomenal characteristics. When the differentiation results from chromatic gradients alone, a distant, indefinite figure is reported, and modifications in the appearance of the fog are slight. Those situations involving dominant short wavelength seem more indefinite and distant than those involving dominant long wavelength. The addition of an intensive gradient to the chromatic gradient increases definiteness of the figure-ground segregation.

The effect of moderate differences in intensity is to produce a relatively definite separation of figure and ground. The density of fog is considerably reduced. Increased differences in intensity further separate figure and ground. The fog disappears. The figure is seen as closer than the ground. Their modes of appearance differ and a distinct boundary separates them. Those factors which increase the apparent distance of the field consistently decrease the apparent distance of the spot and the density of the fog.

The phenomenal characteristics of the *Ganzfeld* do not appear to be independently determined. Any change in the stimulus-distribution that modifies one aspect of experience produces concomitant changes in the other phenomenal characteristics of the *Ganzfeld*. The distribution of stimulation rather than the nature of the local stimulus seems to determine what is experienced.

CERTAIN RELATIONS AMONG EXPERIENCED CONTINGENCIES, ASSOCIATIVE STRUC- TURE, AND CONTINGENCIES IN ENCODED MESSAGES

By CHARLES E. OSGOOD, University of Illinois, and LOIS ANDERSON,
Massachusetts Institute of Technology

This experiment is a test of some of the assumptions and implications of association-theory; namely (1) that the associative structure of a set of meaningful items depends upon the past experience of S ; (2) that patterning in the spontaneous sequential output of these items by S depends directly upon the associative structure and indirectly upon his experience; and (3) that the strengths of both associative and sequential probabilities depend upon the relative frequencies with which items have paired in experience. In the experiment to be reported, a hierarchical structure of input-pairings of each item with other items was imposed with a view to testing these hypotheses. The implications of the results for a method of content-analysis known as *contingency content-analysis* will be considered.

HYPOTHESES AND DESIGN

Our general assumption is that (a) *contingencies in experience* of varied frequency come to be represented in (b) an *individual's associative structure* by patterns of association of varying strengths, which help determine (c) *the measured contingencies in messages* produced by this individual. We require a simple situation in which we can measure $F_a(b) > F_a(c) > F_a(d) > \dots F_a(n)$, the varying frequencies in experience with which a is followed by $b, c, d, \dots n$; $P_a(b), P_a(c), P_a(d), \dots P_a(n)$, the varying probabilities with which S s exposed to the above experience will associate items $b, c, d, \dots n$ when *some other person* (E) gives a —this provides a measure of associative structure (associative probability); and $P_a^*(b), P_a^*(c), P_a^*(d), \dots P_a^*(n)$, the varying probabilities with which S s with above experience will produce items $b, c, d, \dots n$ after *they themselves* have produced a —this provides a measure of message-contingency (transitional probability).

If we think of the S in this experiment as a communicating unit in the sense of information theory, $F_a(b)$ is the input to the unit and $P_a^*(b)$ is the output. The input is so determined by E that $F_a(b) > F_a(c) > F_a(d)$, e.g. a hierarchical contingency-structure. It is assumed that previous to this experience the associations among these items are random across S s, and materials for the experiment are selected to approximate this condition.

Hypothesis I. Exposure to a sequence of paired items such that $F_a(b) > F_a(c) > F_a(d)$ will result in a non-chance associative structure among these items such that $P_a(b) > P_a(c) > P_a(d)$. This dependency-relation between input and associative

* Received for publication October 29, 1956.

structure may be demonstrated in several ways, including correlation.

Hypothesis II. Given an associative structure such that $P_a(b) > P_a(c) > P_a(d)$ in a set of Ss, sequential messages by these Ss limited to these items will display contingencies (transitional probabilities) such that $P_a^*(b) > P_a^*(c) > P_a^*(d)$. This dependency-relation between associative structure and output also can be demonstrated by correlation and by other measures.

Hypothesis III. Given exposure to a sequence of paired items such that $F_a(b) > F_a(c) > F_a(d)$ and subsequent production of sequential messages limited to these items, message-contingencies will be such that $P_a^*(b) > P_a^*(c) > P_a^*(d)$. This dependency-relation between input and output assumes mediation via S's associative structure and can be measured in similar ways.

Hypothesis IV. The dependency-relation between associative structure and message-contingency will be greater than that between input-contingency and message-contingency. This hypothesis derives from the assumption that message-contingencies depend directly upon S's associative structure and only mediatedly upon his experience. To the extent that individual Ss have prior associative experience with the items, these associations will also influence the final structure. In the design of the experiment, evaluation of this hypothesis requires that our two tests be given in alternate orders (e.g. Group I: input—association-test—message-production; Group II: input—message-production—association-test); otherwise positive results for this hypothesis could be due simply to recency-effects.

Hypothesis V. The degree of dependence (a) of associative structure upon input-contingency and (b) of message-contingency upon input-contingency will be a direct function of input-contingency, $F_a(b)$. In other words, we assume that modification of associative structure (and hence of transitional message-structure) varies with the frequency with which events are paired in experience—a straightforward principle of association. This implies that the more frequent pairings have been in experience, the more significant will be the deviations of associative and transitional probabilities from chance.

Hypothesis VI. The degree of dependence between associative structure and message-contingency will be relatively independent of input-contingency. This assumes that whatever preexperimental associations between items exist in individual Ss will also determine both associative and transitional (message) contingencies, hence dependency-relations here should be relatively unaffected by experimental inputs. In other words, although the items associated will change as a function of specific input, the degree of relation will be relatively constant as a function of $F_a(b)$.

PROCEDURE

Materials. Ten common names for girls were selected, each beginning with a different letter, and none rhyming with any other. Care was taken to avoid including any pairs of names for which there was known cultural association.¹ In the materials presented to the Ss, each name was paired with one other in the list 6 times, with a second name 3 times, and with a third name once; the design was such that in the total 100 paired-associates, every name appeared 10 times as the first member

¹ A pilot-study was conducted with nonsense-syllables of low association-value. The effect of 100 pairings proved, however, to be stultifying.

of the pair and 10 times as the second member of the pair—thus sheer frequency and position were constant for all items. In the 100 pairings, each name appeared with 6 other names at various frequencies and in different relations (forward or backward), but never with 3 of the other names; e.g. for the item Josephine: Josephine—Beatrice, 6; Louise—Josephine, 6; Josephine—Cynthia, 3; Gladys—Josephine, 3; Josephine—Hazel, 1; Esther—Josephine, 1; with Sarah, Isabelle, or Valerie, 0.

Two groups of 100 Ss each were shown 100 frames of a single-frame projection film-strip successively, with 3 sec. between successive frames. Each frame showed a pair of names, e.g. Beatrice—Louise. The order of frames with respect to pairings was random, with the restriction that the same name should not be repeated within any set of 3 frames. The Ss were told simply to familiarize themselves as much as possible with the names which had been paired randomly for convenience in presentation. Judging from Es' observations, interest in the names was sufficient to maintain an adequate degree of attention throughout the task. Following presentation of the 100 frames, two different tests were made, in reverse order, for the two groups:

Associative test. Each of the 10 names was shown on the screen separately, 8 sec. apart, and the Ss were instructed to write the first other girl's name that occurred to them. In this test, *E* provides the stimulus and *S* gives a response.

Transitional test. The Ss were given little booklets and instructed to write one girl's name on each page, turning the page each time and not looking back, for as many pages as they could fill in. In this test, the stimulus for each response derives from *S*'s previous behavior, i.e. this is emitted rather than elicited behavior.

The group that had the associative test first and the transitional test second will be referred to as Group I; the other group will be called Group II.

Treatment of the data. The two different tests were first 'scored' for individual Ss. For the associative test, it was simply recorded if each response was one which had been paired with the stimulus-name in the input, with what frequency it had been paired, and in what direction (forward association or backward association in terms of location on the frame). In the transitional test, only the first time that a given name appeared was counted as a stimulus; if this name appeared again in the booklet, it was not used. Three tables were then formed for each group of Ss:

Input-table. This table gives, for each name, the relative frequency (%) with which it had been paired with every other name in the input, without regard to direction. Thus, an item paired 6 times with another would account for 30% of its appearances; with another 3 times, 15% of its appearances; and with another once, 5% of its appearances—50% of these pairings are forward and 50% are backward. There are therefore two 30% pairings, two 15% pairings and two 5% pairings for each item. These input-values are given in Table IA.

Associative table. This table shows, for each name, the relative frequency (%) of Ss giving each of the 10 possible responses. Thus, if 16 of the 100 Ss in Group I wrote 'Beatrice' when they saw 'Josephine' on the screen, 16% was entered in the appropriate cell. These associative values are given as Tables IB and IIB for Groups I and II, respectively.

Transitional table. This table shows, for each self-produced stimulus-name, the relative frequency (%) of Ss giving each of the 10 possible response-names. Thus, if the name 'Josephine' appeared in the booklets of 79 Ss in Group I and was followed

immediately by the name 'Hazel' in the booklets of 11 of these Ss, 14% was entered in the appropriate cell. These transitional values are given as Tables IC and IIC for Groups I and II, respectively.

We wish to demonstrate that the relative frequencies of both associative and transitional pairings have been influenced by the imposed input-pairing, as well as to show that the degree of effect is different in the two cases. It will be noted, however, that the input-tables are made up of discrete, stepwise entries (30%, 15%, 5%,

TABLE I
BASIC DATA FOR GROUP I

	Stimulus-name	Jos.	Bea.	Est.	Sar.	Isa.	Lou.	Haz.	Val.	Cyn.	Gla.	r_{AB}	r_{AC}	r_{BC}
A. Input	Josephine	00	30	05	00	00	30	05	00	15	15			
	Beatrice	30	00	00	05	15	05	15	00	30	00			
	Esther	05	00	00	15	30	15	05	00	00	00			
	Sarah	00	05	15	00	00	00	30	30	15	05			
	Isabelle	00	15	30	00	00	15	00	30	05	05			
	Louise	30	05	15	00	15	00	00	05	00	30			
	Hazel	05	15	05	30	00	00	00	15	30	00			
	Valerie	00	00	00	30	30	05	15	00	05	15			
	Cynthia	15	30	00	15	05	00	30	05	00	00			
	Gladys	15	00	30	05	05	30	00	15	00	00			
B. Associative response (first)	Josephine	00	16	02	08	18	17	19	05	07	08	0.23		
	Beatrice	11	00	04	09	09	07	16	07	31	06	0.82		
	Esther	11	03	01	14	11	11	17	07	04	20	0.64		
	Sarah	09	04	15	01	07	07	26	16	10	04	0.62		
	Isabelle	27	11	12	12	00	10	06	10	05	09	0.21		
	Louise	18	06	14	14	01	06	05	09	13	03	0.45		
	Hazel	06	07	10	39	03	08	00	16	09	05	0.65		
	Valerie	18	10	05	15	08	14	11	02	07	08	0.23		
	Cynthia	06	32	04	11	05	10	11	04	01	05	0.46		
C. Transitional response (second)	Gladys	13	04	16	12	06	02	14	09	06	01	0.63		
												$r = 0.494$		
	Josephine	00	06	09	09	13	19	14	11	07	13		0.10	0.50
	Beatrice	11	01	04	06	04	08	11	06	37	11		0.50	0.57
	Esther	11	11	00	09	07	14	13	10	11	14		0.32	0.38
	Sarah	13	04	10	00	10	14	10	25	10	06		0.19	0.54
	Isabelle	27	03	09	10	00	12	07	12	15	06		-0.05	0.25
	Louise	10	08	06	19	09	00	10	15	06	17		0.21	0.21
	Hazel	17	04	10	16	20	05	00	15	07	05		0.15	0.31
	Valerie	18	12	14	05	11	07	12	01	11	10		-0.31	0.09
	Cynthia	08	12	11	09	06	10	23	09	00	13		0.18	0.50
	Gladys	06	10	18	14	09	11	18	09	05	00		0.26	0.50
												$\bar{r} = 0.155$	$\bar{r} = 0.385$	

or 0%); furthermore, there is probably some lack of independence among row-values in the associative and transitional tables, in that S's use of one name to a prior name would decrease the likelihood of his using the same name to a subsequent name. We have therefore used Kendall's *Tau*, making adjustments for tied ranks where they appeared, and applied it separately to each row.² Because of possible lack of independence among rows, these rank-order correlations cannot be cumulated in assessing significance, but their general levels of magnitude for the various conditions will permit us to draw the relevant conclusions. Certain additional treatments of the data will be described in connection with specific results.

RESULTS

Kendall's *Tau* values for relations between input-contingency and asso-

² M. G. Kendall, *Rank Correlation Methods*, 1948, 1-36. The writers thank Professor William E. Kappauf of the University of Illinois for suggesting this measure.

ciative structure (r_{AB}), between input and transitional contingency (r_{AO}), and between associative structure and transitional contingency (r_{BO}) are given as indicated in Tables I and II.

With regard to Hypothesis I, it can be seen that the r s between input-frequency and associative probability (r_{AB}) are consistently positive for all rows in both groups, as predicted. The fact that the average r for Group

TABLE II
BASIC DATA FOR GROUP II

	Stimulus-names	Jos.	Bea.	Est.	Sar.	Isa.	Lou.	Haz.	Val.	Cyn.	Gla.	r_{AB}	r_{AC}	r_{BC}
B. Associative response (second)	Josephine	01	06	06	11	09	13	21	08	13	02	0.33		
	Beatrice	13	00	05	13	05	07	22	04	26	02	0.74		
	Esther	12	05	00	09	09	16	15	07	10	13	0.43		
	Sarah	11	06	14	02	04	09	14	17	11	09	0.74		
	Isabelle	24	15	11	08	00	08	08	06	15	03	0.03		
	Louise	18	08	15	10	13	00	13	10	05	06	0.43		
	Hazel	09	16	08	20	08	19	00	07	07	05	0.26		
	Valerie	15	10	07	05	15	13	15	00	09	06	0.10		
	Cynthia	03	30	09	06	09	09	16	10	00	05	0.41		
	Gladys	05	05	09	08	08	16	18	08	16	00	0.18		
												$\bar{r}=0.365$		
C. Transitional response (first)	Josephine	00	11	06	08	11	17	11	07	04	07		0.26	0.51
	Beatrice	08	00	03	05	04	04	10	09	28	10		0.36	0.44
	Esther	06	04	00	05	04	07	12	05	02	08		0.41	0.64
	Sarah	11	05	05	02	09	06	06	14	04	02		0.10	0.35
	Isabelle	10	08	04	08	00	07	05	08	14	04		0.03	0.59
	Louise	08	08	02	10	02	01	09	08	08	06		-0.30	0.15
	Hazel	12	12	07	14	12	04	01	11	11	06		0.53	0.50
	Valerie	15	09	07	06	08	10	06	00	08	06		-0.08	0.62
	Cynthia	11	06	12	10	09	03	21	10	00	08		0.40	0.21
	Gladys	03	07	07	04	06	06	15	06	09	00		-0.16	0.63
												$\bar{r}=0.155$	$\bar{r}=0.464$	

I (0.494) is somewhat higher than that for Group II (0.365) presumably reflects the interpolation of the transitional test in the latter group. The correlations between input-contingencies and transitional contingencies (r_{AC}) are inconsistent as to sign, although their average for both groups (0.155) is in the predicted direction. There is therefore little support for Hypothesis III.

With regard to Hypothesis II, it may be seen that the r s between associative and transitional probabilities (r_{BO}) are consistently positive for both groups, as predicted. Furthermore, comparison of the r_{BO} -arrays for Groups I and II indicates that they are not significantly different in magnitude ($I > II$ in 6 cases, $I < II$ in 4 cases), despite the tendency for associative structure to be more dependent upon input in Group I. This outcome is consistent with Hypothesis VI. Finally, by comparing the arrays of r_{AO} with those for r_{BO} within each group, it can be seen that the latter, with only a few exceptions, are greater in magnitude than the former. In other words, as anticipated from Hypothesis IV, transitional contingencies in the output are more dependent upon associative structure than upon the original input-pairings.

To test Hypothesis V—that the degree of dependency of both associative and transitional probability upon the input varies with the absolute frequency of input-pairing—it is necessary to demonstrate that the predictability of response-name from knowing the stimulus-name varies with the frequency of pairing in the input, *e.g.* that correct association for 6 pairings is greater than for 3 pairings, $3 > 1$, and so forth.

For each *S* the number of correct responses given in each frequency-

TABLE III
PERCENTAGES OF *Ss* GIVING CORRECT RESPONSES IN VARIOUS
FREQUENCY-CATEGORIES
(Averages over 10 stimulus-words)

No. of pairings	Group I				Group II			
	Associative		Transitional		Associative		Transitional	
	Forw'd.	Backw'd.	Forw'd.	Backw'd.	Forw'd.	Backw'd.	Forw'd.	Backw'd.
6	19	16	16	12	17	12	16	11
3	12	10	09	09	12	10	09	09
1	09	07	09	10	12	08	11	10

SIGNIFICANCE OF DIFFERENCES BETWEEN VARIOUS FREQUENCY-CATEGORIES
(DIXON-MOOD TEST)

	Group I				Group II			
	Associative		Transitional		Associative		Transitional	
Forward	6>3	1%	6>3	1%	6>3	1%	6>3	5%
	6>1	1%	6>1	10%	6>1	5%	6>1	5%
	3>1		3>1		3>1		3>1	
Backward	6>3	5%	6>3		6>3	10%	6>3	
	6>1	1%	6>1		6>1	5%	6>1	
	3>1	10%	3>1		3>1		3>1	
Six pairings	F>B	10%	F>B	11%	F>B	5%	F>B	10%

category was recorded. A Dixon-Mood sign-test was used to determine significance, *e.g.* for how many *Ss* the frequency of correct 6-responses was greater than the frequency of correct 3-responses, and so on. Both forward and backward directions of association were analyzed in this manner. Table III shows the percentage of *Ss* giving correct responses in the various categories (upper) and the significance-levels for the several comparisons (lower). For all conditions except the backward direction on the transitional tests, input-frequencies of 6 pairings yielded significantly more correct responses than either 3 or 1 pairings; this result is consistent with Hypothesis V. In no case, however, is the difference in frequency of

correct responses between 3 and 1 input-levels significant at the 5% level; this indicates that for the conditions of this experiment a pairing-frequency of 3 does not reach the threshold for effective response-evocation (values of neither 1 nor 3 are significantly above those for 0 pairing-frequency, chance). Finally, taking only the 6-pairing level, where significant effects are found, the last entries in Table III indicate that for both associative and transitional tests in both groups the forward direction of association is somewhat superior to the backward direction (at least 11% significance-level or better). Given the standard reading habits in our culture, the left-hand name on each frame was probably seen first, and many studies of association have shown that forward connections tend to be stronger than backward.³

As a further check on the assumption that relations between associative structure and transitional structure are relatively independent of immediate input-experience (Hypothesis VI), two extreme subgroups were formed from each of Groups I and II on the basis of correspondence between input-contingencies and transitional output: (1) 'high'-correspondence subgroups (defined as Ss giving 3 or more pairs in the transitional test as they were paired 6 times in the input); (2) 'low'-correspondence subgroups (defined as Ss giving no pairs in the transitional test as they were paired 6 times in the input). The question is this: will there be any differences between these extreme groups in the degree to which responses in the transitional test are dependent on (predictable from) responses to the same stimuli on the associative test? If not, then we are justified in concluding that transitional (message) contingencies reflect associative structure but not necessarily any particular input-experience.

For each subgroup a matrix was formed in which the rows represented the response-names given to a stimulus-name on the associative test and the columns represented the response-names given to the same stimulus-names on the transitional test. All values falling in the diagonal cells of these matrices therefore represent 'hits', *i.e.* cases in which the same response to a stimulus-name was made on both tests. For both Groups I and II, and for both 'high' and 'low', the frequencies of 'hits' were significantly greater than would be expected on a chance basis (at the 1% level or better). In general, then, these data are consistent with Hypothesis VI. That there was nevertheless a slightly greater degree of correspondence

³ J. J. Jenkins, W. A. Russell, *et. al.*, Technical report No. 1 and summary report, 1955, under ONR Contract N8onr-66216; Jenkins and Russell, Associative clustering during recall, *J. abn. soc. Psychol.*, 47, 1952, 818-821.

for the 'highs' is suggested by the percentages of total responses falling in the diagonal cells for these groups: 21% for Group I 'highs,' 24% for Group II 'highs,' 15% for Group I 'lows,' and 18% for Group II 'lows.'

DISCUSSION

From the point of view of associational-theory in psychology, these findings support the basic assumption that a hierarchical system of pairings of input-events comes to be reflected in a corresponding hierarchical associative structure; the greater the impressed frequency of input pairing, the greater the correspondence. It also supports an assumption usually made but seldom directly tested, that spontaneously emitted sequential items will reflect the associative structure of the emitting S.⁴ The fact that the spontaneously emitted transitional sequences are more closely (though not perfectly) correlated with the independently measured associative bonds than they are with the experimentally manipulated input-frequencies—and this being equally true for Group II where the transitional tests came first—is evidence for a two-stage (at least) mediating system between input- and output-contingencies.

This experiment also was designed as a test of some of the assumptions necessary for *contingency content analysis*, in which the analyst utilizes the greater-than or less-than chance contingencies among items in the messages produced by some source as a basis for making inferences about the associative structure of that source and, more remotely, about the life-experiences of that source.⁵ On general psychological grounds, it seems reasonable to assume that greater-than-chance contingencies among items in messages are indicative of associations among these items in the thinking of the source. If, in the experience of the source, 'Food supply' and 'Occupied countries' often have occurred together (either as events or as signs of events), the occurrence of one of them should become a condition for subsequent occurrence of the other, and hence the writing about or speaking (*i.e.* 'encoding') of one should tend to elicit production of the other. It also seems reasonable to assume that *less-than-chance* contingency among items in messages would be indicative of dissociations in the thinking of the source. If, in the experience of a patient in psychotherapy, for example,

⁴ Cf. W. A. Bousfield and B. H. Cohen, Technical reports No. 4 and 5 under ONR Contract Nonr-621 (00); Jenkins and Russell, *J. abn. soc. Psychol.*, 47, 1952, 818-821; Jenkins and Russell, Technical reports No. 2 and 3 under ONR Contract N8onr-66216.

⁵ A. L. Baldwin, Personal structure analysis: A statistical method for investigating the single personality, *J. abn. soc. Psychol.*, 37, 1942, 163-183; see also the report of the SSRC Conference on Content Analysis, in preparation.

'Mother' and 'Heterosexual experience' often have been associated—but under conditions of fear and anxiety—the occurrence of one of them might lead to inhibition of the other. The present experiment, however, dealt only with degrees of association.

In applying contingency content-analysis to real problems, such as propaganda and psychotherapy, we seldom if ever have any criteria against which to validate our inferences. Usually we have only the messages produced, not the source who produced them (who could give us other, independent indices of his associative structure) and certainly not his past history. To test the basic assumptions of this method, therefore, it is necessary to devise a controlled experimental situation in which (a) the experiential, input-history of the source can be known approximately, and (b) the associative structure can be estimated independently of the message-contingencies. The present experiment provided such conditions.

In terms of the practical problems of contingency content-analysis, our results imply the following state of affairs: The associative structure (relations among ideas) of the source of a message being analyzed depends upon the experience of this individual, and hence experiential life-history is to some extent inferable from associative structure (Hypothesis I). The measured contingencies between events in messages produced by the source are indicative of the patterns of association of ideas in this source, and hence one may base inferences about the associative pattern of the source upon these measured contingencies (Hypothesis II). Given the mediation by the associative structure of the source, measured contingencies in messages can be used to make inferences about the life-history of the source (Hypothesis III). Since, however, measured contingencies in messages are only indirectly related to any particular set of experiences of the source, via the effects of such experiences upon modifying associative structure, inferences about specific life-experiences from message-contingencies are more liable to error (Hypothesis IV). On the other hand, the more repetitive certain experienced contingencies in the life-history of the source, the more likely is this experience to influence associative structure and hence message-contingency (Hypothesis V). Regardless of the patternings of events in life-history, however, measured contingencies in messages will provide valid evidence for the associative structure of the source (Hypothesis VI).

The burden of this study is that contingency content-analysis can provide a valid index of the associative patterns of the source, but only a mediate and tenuous index of his life-history. Two limitations, however, must be mentioned: In the first place, the materials and the setting of this

experiment are rather remote from the typical materials and settings of ordinary content-analysis. Secondly, it is probably obvious that the assumptions made here, and hence the method, apply best to informal messages, such as diaries, letters, spontaneous speech, and the like; the more carefully and deliberately planned a message, as for propaganda and educational purposes, the more remote, presumably, becomes the connection between message-contingencies and the 'real' associative structure of the source.

SUMMARY

This experiment was designed to test certain assumptions about the formation and utilization of associations that underlie a contingency-method of content-analysis; namely, (1) that the associative structure of a source depends upon the contingencies among events in his life-experience, and (2) that inferences as to this associative structure can be made from the contingencies among items in messages he produces. Conditions in which the contingencies among input-events to human 'sources' could be manipulated at least in part, and in which the resultant associative structures could be determined independently of the contingencies among the items in the 'messages' (transitional outputs) produced, were arranged. The results show that associative structure depends upon input-contingency, and output-contingency depends upon associative structure, both to a significant degree. Output-contingencies in 'messages' are, however, only remotely related to specific experimental input-contingencies, *i.e.* non-chance associations among items existed in individual Ss prior to the experimental treatment and also influenced the output ('message') contingencies. In general, the degree to which input influences both associative structure and finally encoded 'message'-structure is a function of the frequency of input-events.

EFFECT OF SHOCK-CESSATION AS AN INCIDENTAL REWARD IN VERBAL LEARNING

By LYMAN W. PORTER, University of California

Experiments on incidental learning have yielded contradictory results with respect to Thorndike's hypothesis that rewards act directly and automatically to establish and strengthen associative bonds.¹ Another explanation that has been advanced to explain the effect of reward (*E*'s saying 'right') in these studies is that the reward produces implicit rehearsal of the rewarded responses even though the subject (*S*) has not been instructed to learn. Although the results of one of these recent experiments designed to test the rehearsal explanation failed to confirm it,² none of the experiments has entirely eliminated the possibility of rehearsal. If reward does produce differential rehearsal, the instances of learning which have been reported could be explained in terms of contiguity and frequency without recourse to Thorndike's principle. What is needed is for an unambiguous reward to be presented under conditions comparable to those of the above studies, but in which it is extremely unlikely that the reward would cause *S* to give differential review to the rewarded associations. The present experiment proposes to do this.

Method. As in the work cited, *S* was asked to guess numbers from 1 to 10 in response to a series of stimulus-words. In the previous experiments, certain of the responses were called 'right' by *E*, and others were called 'wrong.' *S* was informed that the experiment was concerned with extrasensory perception (ESP), and that responses correct on one trial might or might not be correct on a subsequent trial. This was the procedure devised by Wallach and Henle to provide reward in the absence of intention to learn.

The innovation in the present experiment was the use of a different "satisfier" which would not appear to *S* as in any sense contingent upon the particular response he made. *E* did not say 'right' or 'wrong'; instead, the "satisfier" was cessation of 24-sec. electric shock adjusted to be as strong as *S* could tolerate. Shock-cessation was chosen primarily because it is generally recognized as a rewarding state of affairs in those present-day learning theories that utilize Thorndikian direct-action hypothesis. The basic plan of the experiment was to tell *S* that the effect of distraction on ESP was being investigated, and that the shock would be 'on' part of the time and 'off' part of the time according to a predetermined schedule. Shock-cessation could thus be correlated with certain of *S*'s responses without leading *S* to believe that this was a learning experiment, or that any particular response of his caused the termination of shock.

* Received for publication September 12, 1956. This paper is adapted from a doctoral dissertation submitted to Yale University. The writer is indebted to Drs. Fred D. Sheffield, Neal E. Miller, and William Kessen for assistance on this study.

¹ Hans Wallach and Mary Henle, An experimental analysis of the law of effect, *J. exp. Psychol.*, 28, 1941, 340-349; A further study of the function of reward, *ibid.*, 30, 1942, 147-160; Leo Postman and P. A. Adams, Performance variables in the experimental analysis of the law of effect, this JOURNAL, 67, 1954, 612-631; 'Isolation' and the law of effect, this JOURNAL, 68, 1955, 96-105; M. E. Bitterman, Information and effect in incidental learning, this JOURNAL, 69, 1956, 410-416; L. W. Porter, The effect of "right" in a modified Thorndikian situation, this JOURNAL, 70, 1957, 219-226.

² Porter, 219-226.

There should under such circumstances, be no tendency for differential implicit rehearsal of rewarded responses.

Subjects. The Ss were 64 Yale students obtained through the student employment agency. They were told that they would participate in a psychological experiment but were not given any further details until they came to the laboratory.

Design. A previously used procedure was followed as closely as possible within the limitations imposed by the use of the shock.³ The stimulus-words were the same 28 three-letter words, but only a 6-sec. rate of presentation was used. As part of the balancing procedure, Positions 5-24 in the list were designated beforehand as the 'key' positions to be analyzed. The first time through the list every S received the shock for three periods. The shock always came on immediately after a response, the only delay being the reaction-time of E in pressing a telegraph key, which presumably was about as long as it took E to say 'right' or 'wrong' in the previous experiment. The shock remained on for the next four responses. It was turned off immediately following the fourth response and remained off for the next four responses, after which it was turned on again.

Eight sequences of the on-and-off pattern of shock were used to achieve balancing for the 20 key positions. These eight sequences involved the same pattern of alternating sets of four shock-on and four shock-off words, and differed only in the starting point, which ranged from the first to the eighth position. Using eight sequences meant that, for each replication of eight Ss, responses in Positions 5-24 would be used once each in each of the eight different places in the shock-pattern.

The 64 Ss were divided into two groups of 32 each. One group had the identical shock-pattern on the first and the second trials with the list. The other group had this shock-pattern on the first trial, but received no shock at all on the second trial. These two groups will be referred to as the shock-shock (*Sb-Sb*) group and the shock-no-shock (*Sb-NSb*) group, respectively. The reason for the *Sb-Sb* condition was to have a group in which the cue-patterns (and the pain-drive) remained constant from one trial to the next; otherwise, all Ss would have been tried under the *Sb-NSb* conditions, which were most comparable to those of the previous experiment in which no rewards were given on the second trial.

Apparatus and procedure. The shock used was produced by a General Electric sine-wave generator set at 20 cycles. The intensity varied according to S's tolerance (but remained constant for any one S), and for the group ranged from 10 to 32 v. across S's electrodes. The electrodes consisted of circular pads of saline-soaked cotton contained inside rubber suction-cups, 1 in. in diameter. They were strapped on the volar surface of S's wrist, separated from each other by a distance of $\frac{1}{4}$ in. The generator was on throughout each session, and the connection to S was made by depressing a key.

When S entered the experimental room he was asked to fill out a brief questionnaire about his belief in ESP and then given instructions to read to himself. These instructions informed him that E was investigating the influence on ESP of several factors, such as knowledge of results, rate of presentation of stimuli, and presence of distraction. In connection with the variable of presence of distraction, S was told

³ Porter, 219-226.

electric shock was being used because, while many distractors can be ignored, it is difficult to ignore shock; also there would be a chance to get used to the shock before the start of the actual experiment. After *S* had finished reading the instructions, he was reassured that the shock to be used was not dangerous. Under these conditions all *Ss* agreed to go ahead with the experiment.

Adaptation began with the shock at zero-level and gradually increased up to a point at which *S* reported that it was starting to become painful. It was then turned off for about 15 sec., and *E* asked *S* if he could stand it somewhat higher since "we want to get a level as high as you think you can stand for 30 sec." The shock was then turned on again at the level at which it was previously turned off and gradually increased from that level. When *S* reported he had reached his limit, the shock was turned off again. Before turning on the shock for the third time, *E* said he wished to turn the shock on at the last level and keep it on for 30 sec., if it could be tolerated. The shock was then given for 30 sec., voltage and amperage levels being recorded at the end of this period.

The *S* was next given further instructions to read, which stated that: (1) his task was to attempt to select correctly the number from 1 to 10 that went with each word; (2) he was to give his response as quickly as possible after each word was presented; (3) the list of words would be presented "more than once," and a "number which is correct on the first presentation *may* or *may not* be correct on the next;" (4) the shock would come on more than once and would remain on for "predetermined portions of the list (but for no longer than 30 sec. at a time);" and (5) he would "not be able to ignore the shock, of course," but he should not pay any particular attention to it since his task was to "attempt to select the correct numbers that go with the words."

After *S* had read these instructions, he was told that he would get a brief practice period of responding at the rate the machine would present the words in the actual experiment. *S* was then given practice at calling out a number to each of a series of four dashes at the experimental rate. At the conclusion of this practice he was told that the experiment would now begin and that he was to "try to use *ESP* to select the correct numbers from 1 to 10 that go with the words." He was told that after the experiment was over he would be informed about how many "hits" he had made. *S* was then given the first list, the shock being administered at the appropriate intervals. At the conclusion of the first list, *E* told *S*: "Now we'll go through the words again. Remember, give the first number that comes to mind for each word." The *Ss* who did not receive shock the second time were told that there would be no shock and the electrodes were removed before continuing. At the conclusion of the second list, *E* told *S* the actual number of *ESP* 'hits' he had made on each of his trials, and then interviewed *S* with a standard set of questions designed to obtain *S*'s comments about the experiment (especially the shock) without giving away the true purpose. After questioning *S*, *E* briefly discussed the ostensible 'purpose' and asked him not to discuss the experiment with anyone else.

Results. The interviews after the experiment confirmed the expectation that reduction of the shock used could properly be classed as a 'satisfier,' and that the *Ss* were naïve as to the purpose of the experiment. The interviews revealed that the shock was painful or definitely annoying (81% of *Ss*); that it was a relief to have it turned off (92%); and that it felt as strong or stronger at the end of the last

shock as it did at the beginning of the first shock (83%). Only six Ss reported special hypotheses involving belief that their behavior influenced the shock. Of these six, only two thought their behavior might have influenced the *offset* of the shock.

The results on repetition do not show any tendency for the 'reward' to increase the repetition of preceding word-number pairs. The percentages of repetition of responses for each of the eight categories of presence and absence of shock in the *Sb-Sb* and *Sb-NSb* groups are shown in Table I. The most important comparison to be made from the data is that between percentage of repetition of the shock-offset

TABLE I
PERCENTAGE OF REPETITION AT VARIOUS POSITIONS IN THE SHOCKING-PATTERN

		Positions					
	Group	1	2	3	1-3	4*	1-4
Shock-off	<i>Sh-Sh</i>	21.3	25.0	19.7	22.0	12.4	19.6
	<i>Sh-NSh</i>	14.6	16.6	23.9	18.4	16.1	17.8
	Total	18.0	20.8	21.8	20.2	14.2	18.7
Shock-on	<i>Sh-Sh</i>	22.3	6.8	9.9	13.0	12.9	13.0
	<i>Sh-NSh</i>	19.2	21.8	23.4	21.4	14.0	19.6
	Total	20.8	14.3	16.6	17.2	13.5	16.3

* Responses in this position were followed by onset of shock for shock-off words, and by offset of shock for shock-on words.

responses and the average percentage of repetition for the preceding three responses. This comparison is analogous to that used in previous studies. Shock-offset in the present experiment was analogous to saying 'right' in the previous experiments and leaving shock on was analogous to saying 'wrong.' The reason for using only the three shock-on responses preceding offset is to hold constant the presence of shock in both 'rewarded' and 'punished' responses. The data for this comparison are presented in Table I. The table shows that, in both the *Sb-Sb* and *Sb-NSb* groups, the shock-offset response was, if anything, repeated less often than the average of the three preceding responses. These results therefore fail to support the direct-action hypothesis. A parallel analysis can be made for shock-onset responses, using as a basis for comparison the three preceding responses in which the shock was off. For both groups, there was a lower level of repetition of responses immediately followed by onset of shock. For the *Sb-Sb* group this difference was significant beyond the 1% level of confidence.

The foregoing results suggest that the effect of the sudden change in the environment—whether onset or offset of shock—was to decrease the level of repetition of responses just preceding the change. If the onset and offset conditions are combined and compared with the other conditions combined, decreased repetition is found to be associated with environmental change in both the *Sb-Sb* and *Sb-NSb* groups. The difference in each group approached significance and was beyond the 5% level for the two groups combined.

It might be argued that the effect of the shock-cessation might show up, not on the response to the same word the second time through the list, but rather on the first response the next time the shock came on in the same list. The results did not confirm such an argument: the reinforced number was not given more often at the beginning of the next shock-period than a 'control' number (number said just preceding the rewarded number). A related argument might be that whereas shock-cessation did not strengthen the connection of the responses to any particular cue-pattern, it might have increased the 'operant level' of rewarded numbers. In fact, however, the shock-offset number was not given significantly more often during the eight following responses than was a 'control' number (*i.e.* the third shock-on number).

A suprising feature of the data is the low over-all level of repetition of shock-on pairs in the *Sb-Sb* group. In the *Sb-NSb* group, there was only a slight difference between shock-off and shock-on averages (using the averages of Positions 1-4), but there was a larger difference between the averages in the *Sb-Sb* group (significant at the 5% level in a *t*-test of 32 individual difference-scores between shock-on and shock-off conditions). The second-order difference between the *Sb-Sb* and *Sb-NSb* groups with respect to the shock-on versus shock-off conditions was significant at about the 6% level. This difference between the two conditions was even more significant (1%) if only the second and third shock-on responses were compared with the second and third shock-off responses. (The second and third responses in each case might be considered more representative of stable shock-on and shock-off conditions since the responses were neither preceded nor followed by change).

These findings suggest that the presence of shock during acquisition did not affect repetition; rather, the shock interfered only during the test-trial—the second time through the list—in the *Sb-Sb* group. Thus, shock-on responses were not repeated less than shock-off responses in the *Sb-NSb* group, where shock was present only during the first presentation of the list, but they were repeated less in the *Sb-Sb* group, in which shock was present on both trials. It appears, therefore, that all of the effect is on the side of interference during testing.

Discussion and conclusions. The results of this experiment show no evidence that cessation of shock produced a greater frequency of repetitions. The implication from a direct-action interpretation of previous experiments is that shock-cessation should increase repetition in the same way as the saying of 'right' in an *ESP* experiment. This expectation was not confirmed in the present experiment, where conditions motivating differential implicit rehearsal of rewarded responses were presumably absent.

The assumption that cessation of shock would not provide *differential* rehearsal of contiguous responses does not rule out the possibility that some momentary implicit rehearsal of 'just-said' responses might take place. It could be assumed, for example, that the sudden change from no-shock to shock (and vice versa) immediately following a response would interrupt this brief implicit 'rehearsal' of a pair. In such a case, less repetition of that response, the next time the same word appears, would be expected relative to responses followed by stimulus change. This interpretation may be used to explain the tendency for stimulus-change to be associated with low level of repetition. It will be noted that the above interpretation, uses a contiguity

principle of associative learning, the differential effects being explained in terms of differential amounts of practice only.

Although the results of the present experiment tend to rule out an explanation based on the direct action of rewards, they are not interpreted as lending support to the factor of intent to learn stressed by Wallach and Henle. If rewards act indirectly through rehearsal, this practice could take place without regard to *S*'s intention or conscious set to remember.

It should be pointed out that this experiment is not regarded as an ideal test of the drive-reduction theory of reinforcement. The conditions for learning are poor in the present experiment because of the large number of word-number pairs (and consequent interference) and the small number of practice trials (*i.e.* one). The chief import of the present experiment is that it raises questions as to whether previous results offer valid support for the direct-action hypothesis. Results from this experiment, together with those of the previous experiment that failed to provide positive support for a rehearsal hypothesis, leave in doubt the mechanism of the action of rewards in the experiments to date on incidental learning.

AN ANALYSIS OF A FORM OF THE WATER JAR TEST

By JAMES W. FRICK and J. P. GUILFORD, University of Southern California

In connection with a factor-analytic study of the factors of flexibility, in the Aptitudes Project at the University of Southern California,¹ we were considering the inclusion of various tests that might help to throw light on the nature of the factor of *adaptive flexibility*. Since the Water Jar Test has been very commonly regarded as a measure of a trait of rigidity, we decided to include an adaptation of it in our battery for factor analysis.² For a good description of various forms of the Test, the reader is referred to a recent article by Levitt.³ Our own adaptation will be described further on in the present article.⁴

The typical Water Jar Test is not one we would readily choose to use in a factor-analytic study. As Levitt has pointed out, psychometrically it leaves much to be desired. Furthermore, for our convenience it had to be adapted to group testing. We believe that our adaptation preserves the intention of the test, which is to lead the examinee to believe that the problems can be solved by the same stereotyped formula or operations by giving him five similar initial problems, and to see whether he can and does break the 'set' successfully in solving the last five problems, which can be solved either by the same method or by more direct ones.

There were other doubts concerning the events supposed to occur as the examinee takes the test. Does successful solution of the first five 'Einstellung' or stereotyped problems actually induce a set, as assumed? If it does, it can be hypothesized that the more of the stereotyped solutions the examinee achieves, the longer, on the average, it should take him to break the set and desert the stereotyped solution in attempting the last five problems. This is one hypothesis that we decided to investigate.

The problems in the Water Jar Test are somewhat similar to others in a typical arithmetical-reasoning test. It can be hypothesized that the typical examinee takes the problems in the Water Jar Test as they come, one at a time. He does not, in the usual case, notice the similarity of principle among the first five problems or that they are solved by the same formula. He is therefore ready to take problems 6 to 10 as they come, and if under any motivation for economy of effort at all he may well do them avoiding the stereotyped, longer solution. The typical test of arithmetical reasoning usually has its greatest loading on the factor of *general reasoning*. The test sometimes shows small relation to the factor of *logical evaluation*. We might predict some variances in these two factors, therefore, in the Water Jar Test. If breaking of 'set' is involved in any appreciable number of examinees, which

* Received for publication October 22, 1956.

¹ Under Contract N6onr-23810 with the Office of Naval Research, Personnel, and Training.

² A. S. Luchins, Mechanization in problem solving: The effect of *Einstellung*, *Psychol. Monogr.*, 54, 1942 (No. 248), 1 f.

³ E. E. Levitt, The Water-Jar Einstellung Test as a measure of rigidity, *Psychol. Bull.*, 1956, 53, 347-370.

⁴ We have some reassurance regarding the legitimacy of attempting an adaptation in the fact that Levitt assembles evidence that the Water Jar Test seems to function similarly in spite of variations.

helps them to solve more of the critical items (last five) correctly (by the shorter formulas), then the total score in the test should also have a significant loading on the factor of *adaptive flexibility*. Our expectation was that we should find the test a factorially complex one, as most problem-solving tests are, with *adaptive flexibility* as a possible, but not prominent, component. We shall report here some correlations between the Water Jar Test and tests of the three factors mentioned, and present the loadings in those factors for the Water Jar Test.

Description and administration of the test. Table I shows the illustrative item used in the instructions in presenting the Water Jar Test to the examinees. The

TABLE I

SAMPLE PROBLEM USED IN INSTRUCTING THE EXAMINEES

Given: Three empty jars; 1 qt., 4 qt., and 5 qt. Problem: Measure out exactly 2 qt. of water.

Solution	1-qt. jar	4-qt. jar	5-qt. jar
Step a	0	4	0
Step b	1	3	0
Step c	0	3	0
Step d	1	2	0
Step e			
Step f			

capacities of the jars are given at the top of the table with the amount of water to be measured. Below that are spaces for six steps that may be utilized in solving the problem. The instructions state, however, that no specified number of steps must be used. The table illustrates one solution in four steps, but not by either the stereotyped or the short methods required in the test proper.

Some other changes from Luchins' original form should be mentioned.⁵ The five problems are identical with those of Luchins, except that his Item 4, which is also solvable by a more direct method, was changed. The substituted problem is solvable only by the common formula. The second five problems are solvable by either the long method or short methods. We did not include a 'set-breaker' problem among the last five. Thus, the last five are all critical problems. Five such problems should give a much more reliable indication of the genuineness of the extent to which the supposed set persists. Furthermore, those who succeeded with a set-breaker could, in the group-testing situation, go back and change answers to previous critical items. The rigid person could persist in the stereotyped response to the very end.

Instructions. The instructions to the test were:

In this test you will be given a number of problems in which you are to measure out a required amount of water, using only the containers given. You have available an unlimited supply of water. The water must be measured out in the least number of possible steps. Lettered spaces representing steps are given after each problem, and you are to fill in the blanks for your solution. The six steps shown are merely for your convenience in working and do not mean that six steps must be used.

The test was administered, as part of a 60-variable battery, to 207 aviation cadets at Lackland Air Force Base.⁶ For a total score, the first five items were keyed accord-

⁵ *Op. cit.*, 1 f.

⁶ We are very much indebted to Dr. L. G. Humphreys and Dr. J. W. Bowles and their associates for making this testing possible.

ing to solution by the stereotyped or long method and the last five by direct or short methods.

Results. (a) Persistency of set. It seems to be generally assumed that those examinees who solve all the stereotyped problems by the long method have developed a set. In fact, Luchins adopts five correct stereotyped solutions as the criterion for establishment of a set.⁷ It was hypothesized earlier that the more of the stereotyped problems the examinee solves correctly the more likely he would be to fail to shift to a short solution in the critical problems. Table II presents data that bear on this hypothesis. We have fractionated the results, singling out those Ss who earned

TABLE II

NUMBER AND PERCENTAGE OF INDIVIDUAL Ss EARNING DIFFERENT SCORES IN THE FIRST FIVE ITEMS WHO SHIFT TO DIRECT METHODS DURING THE LAST FIVE ITEMS

No. stereotyped problems solved correctly	No. Ss	Shifting to D solutions		Shifting at Item 6	
		No.	%	No.	%
5	87	66	76	57	66
4	43	25	58	13	30
3	27	6	22	5	19
4 or 5	130	91	70	70	54
3 to 5	157	97	62	75	47

scores of 5, 4, and 3 correct solutions for the first five items. Those who obtained scores of 2 or less on the stereotyped items can be regarded as having developed no set. The chances are that they did not have an opportunity to attempt any of the critical items.

There were 87 Ss who solved all five of the stereotyped items correctly, of whom 66 (76%), shifted to short solutions in doing the critical problems. Actually, 57 of the same group (66%), shifted at Item 6, the first opportunity to do so. When two-thirds of those who have done all five of the set-forming items correctly shift immediately at the first opportunity, it does not seem that an impressive number of sets have been induced.

There were 43 Ss who did four of five stereotyped problems correctly. Of these, 25 (58%) shifted at some time to short solutions of critical problems, 13 (30%) doing so at the first opportunity. Of the 27 Ss who obtained a score of 3 on the stereotyped items, only 6 (22%) shifted to short solutions. Thus, the *more* of the stereotyped problems solved, from 3 to 5, the *larger* is the percentage shifting to short solutions on the critical problems, and this also holds true for shifts at the first opportunity, *i.e.* on Item 6. The differences in percentage of shifting individuals are all statistically significant. The *t*-ratio for the comparison of those earning stereotype scores of 5 and 4 is 2.12. The *t*-ratio pertaining to those with stereotype scores of 5 versus 3 is 5.09. The *t*-ratio for comparison of those earning stereotype scores of 4 versus 3 is 2.95. Clearly, there is a positive correlation between ability to solve the first five problems by the stereotyped method and the ability to solve the last five items by the short methods.

There were some solutions of the critical items by the stereotyped method. Of the 489 total solutions of critical items, 125 were by the long method and 364 were by the short methods. This might be interpreted to mean that there was some carry-

⁷ *Op. cit.*

over of method from work on the stereotyped items. We do not know, however, to what extent the same, long solutions of critical items might have occurred without S's having had prior experience with the stereotyped items. Deriving information on this would call for a control experiment, which we believe has rarely been considered or carried out.

(b) *Factorial information.* We have felt justified in factor analyzing a total score on our Water Jar Test for two reasons. Our analysis of the responses to items mentioned above indicated that the two halves of the test have very much in common. We have not attempted to estimate the correlation between the two halves, as such, since we gave the total test with one time limit of 10 min. working time. Although the test was sufficiently easy to give a mean of 5.4, and there were 30 individuals with a perfect score of 10, there was enough speeding to make such a correlation of very questionable value. Many examinees did not reach the second half of the test.

The other justification for factor analysis of the total score was that if there is any variance in *adaptive flexibility* involved arising from any items in the total test, the analysis should show this fact. This was particularly true since we dichotomized the total-score distribution between scores of 5 and 6. Those showing flexibility by shifting from long to short methods would be likely to earn a score of 6 or higher, although not necessarily so if they missed some of the stereotyped items. Also left out of account are those not reaching Item 6 and who, had there been more time, might have been among the flexible ones. The correlation between the total score and the dichotomy of shift vs. non-shift was 0.87. This is of course inflated by the condition of speeding of the test. It does help, however, to show that the ability to shift is also indicated by the total score. This same conclusion is supported by a correlation of 0.46 between the shift vs. non-shift dichotomy and the score on the first five items only.

The correlation of the total score with three factors—*general reasoning*, *logical evaluation*, and *adaptive flexibility*—are given in Table III. A more complete description of the factorial composition of the test total score will be reported in another place.⁸ Correlations of this score with tests of these three factors are also given in Table III. The predicted relationships of the total score to the first two factors were fulfilled. The small loading for the score in *adaptive flexibility* was not surprising.

It is possible that if scores on the first and second halves of the test were analyzed separately, we should find a higher *adaptive-flexibility* loading for the second part and a lower one for the first part. There would be some question as to which Ss should be included in such an analysis, whether the total test is given as a time-limit test or as a power test, and a question as to whether the two parts should be separately timed and whether this break in continuity would destroy much of the carry-over effect that is under consideration.

Discussion. We agree with the conclusion of Levitt, that the Water Jar Test is not a good psychometric instrument and consequently not a good experimental instrument, for about the same reasons. The frequency distribution that we find is

⁸ J. P. Guilford, J. W. Frick, P. R. Christensen and P. R. Merrifield, A factor-analytic study of flexibility in thinking, *Rep. psychol. Lab.*, No. 18, Los Angeles: University of Southern Calif., 1957.

trimodal (modes at 5, 8, and 10) and otherwise peculiar. Insights are very probably involved, which means unusual interdependencies within clusters of items. There is lack of experimental control that would narrow the kind of insight down to one, or, if not, would enable us to find out what the various kinds are.

As ordinarily used, the Water Jar Test has amounted essentially to a one-item test, with power to discriminate only between two groups; those who shift and those who do not. Even then, the separation is probably unreliable, and the base from which the shift is made differs from person to person. A much longer test, presenting

TABLE III

CORRELATION OF TOTAL SCORE ON THE WATER JAR TEST WITH THREE COMMON FACTORS AND WITH TESTS THAT HAVE SIGNIFICANT LOADINGS ON THE FACTORS

Factor	Correlation*	Test	Correlation
General reasoning	.42	Ship destination	.47
		Circle square triangle	.40
Logical evaluation	.45	Logical reasoning	.42
		Puzzles	.28
Adaptive flexibility	.18	Insight problems	.39
		Match problems II	.28

* Loading on an orthogonal factor.

a number of different sets and opportunities to shift might be an improvement; but then one runs into the danger of *S*'s learning principles of set and shift. If rigidity is the opposite quality to flexibility, fortunately we have better flexibility tests that can be used to investigate rigidity. We shall also have to consider the probability that there are as many as three adaptive-flexibility factors in the area of thinking.⁹

Conclusions. There is little evidence that the Water Jar Tests, in the form used by us, at least, successfully generates a set of any significant strength in many individuals. For the purposes of testing or of experimenting with rigidity, we should like to bring every *S* to the same level of set formation.

From a factor analysis of the total score, we find that most of its non-chance variance is accounted for by factors common to other problem-solving tasks—*general reasoning* and *logical evaluation*. Very little variance in *adaptive flexibility*, which may be regarded as logically opposite to rigidity, was found.

⁹ J. P. Guilford, The structure of intellect, *Psychol. Bull.*, 1956, 53, 267-293.

A CONFIDENCE SCALE DEFINED IN TERMS OF EXPECTED PERCENTAGES

By JOE K. ADAMS, V.A. Hospital, Palo Alto, California

Psychologists have long had an interest in the confidence with which judgments are made. In 1884 Peirce and Jastrow reported an experiment in lifted-weights in which judgments of confidence were obtained with every discrimination.¹ These authors, with only each other as *Os*, demonstrated two facts which have since been confirmed by many experimenters under a considerable variety of laboratory conditions: (1) judgments of confidence are positively related to the difference between the two stimuli being judged, hence to the accuracy of the discrimination; and (2) judgments made with no confidence sometimes yield a higher proportion of correct responses than expected by chance.² In this latter finding Peirce and Jastrow saw implications far beyond the limits of their study, as they wrote: "The insight of females as well as certain 'telepathic' phenomena may be explained in this way. Such faint sensations ought to be fully studied by the psychologist and assiduously cultivated by every man."³

Despite the interest in this phenomenon as a curiosity of the psychological laboratory, there has been almost no attempt to study its variation with experimental conditions, to relate it to individual differences in other variables, or to manipulate it through training.

One limitation on the study of the discrepancy between confidence and performance has been that it seemed possible to compute a discrepancy at only one point, the point of zero confidence or pure guess. The present paper describes the use of a scale which makes it possible to compute discrepancies at several points, and thus to obtain a 'discrepancy-function.'

The scale can be used in any situation in which it is feasible for *O* to indicate his confidence that a given response is correct (or falls into any other independently specifiable category). *O* is instructed to express his confidence in terms of the percentage of responses, made at that particular level of confidence, that he expects to be correct. After he has made a large number of responses and confidence-judgments, the percentage correct for each point on the confidence scale (which he has used with sufficient frequency) is computed and a discrepancy from the expected percentage is thus obtained. If *O* is using the scale in accordance with its intended meaning, and if his confidence-judgments are realistic in the sense of being appropriate to his performance, then of those responses made with '0' (zero) confidence, none should be correct; of those responses made with confidence '10,' about 10% should be correct, etc. In other words, of those responses made with confidence *p*,

* Received for publication November 22, 1956. This research was conducted while the author held the Thomas Welton Stanford Fellowship in Psychical Research at Stanford University, 1955-56. Acknowledgment is made to Mr. Robert S. Witte for assistance in conducting this experiment.

¹C. S. Peirce and Joseph Jastrow, On small differences of sensation, *Mem. Nat. Acad. Sci.*, 3, 1884, 73-83.

²A recent review of the topic of confidence judgments is given by D. M. Johnson, *The Psychology of Thought and Judgment*, 1955, 368-389. For the second point see J. E. Coover, *Experiments in Psychical Research*, 171-226.

³Peirce and Jastrow, 83.

about $p\%$ should be correct. If the responses are restricted to k possibilities, the confidence scale runs from $100/k$ to 100, instead of from 0 to 100; e.g. in a two-category response-situation the confidence scale runs from 50 to 100.

Several exploratory experiments, two of which are described below, have shown that most college students readily grasp the meaning of the scale, if a complete explanation is given, and that once having grasped the meaning they accept the discrepancies obtained as indicative of the extent to which they evaluate the accuracy of their responses in a realistic manner.

EXPERIMENT I

Procedure. Ten members of a laboratory course in learning and perception served as Os. They were seated at distances varying from 7 to 15 ft. from the screen and at varying angles to it. None of the Os knew the purpose of the experiment.

The method was one commonly used in finding the recognitive thresholds of words. Each of 40 words was exposed consecutively 10 times for about .25 sec., beginning with very low and ending with very high illumination. A 500-w. opaque projector was used with shutter and variac. Exposures began at 65 v. and increased in 5-v. steps to 110 v. The words, from four to six letters in length, were type-written in solid capitals, and included both high and low frequency words (Thorndike-Lorge L counts of from 156 to 883 and from 4 to 62 respectively,⁴ in heavy and light type, and with unpleasant and neutral or pleasant connotation (e.g. 'curse' and 'brief'). The Os were first dark-adapted for 15 min. After each exposure they wrote down the word they thought they saw and gave a confidence-judgment of from 0 to 100, the meaning of the scale having been explained as described above. The Os were asked to limit their judgments to the points 0, 10, 20, . . . , 100%. They knew that the same word was being exposed 10 times consecutively.

Results. The 400 pairs of word-responses and confidence-judgments of each O were analyzed separately.⁵ The word-responses were grouped according to their confidence-judgments and the percentage of correct recognitions was computed for each point on the confidence-scale. A word-response was counted correct only if identical with the word exposed. Table I shows the results for every O. Percentages based on at least 10 judgments are shown in boldface.

Two of the 10 Os gave an appreciable (and statistically significant) proportion of correct recognitions with zero confidence. These two also gave large positive discrepancies at other points on the confidence-scale. This finding suggests that underconfidence at the point of 'pure guess' is only a special case of a tendency to place less confidence in one's reports than is warranted by the actual accuracy of the reports.

Five of the remaining Os showed large positive discrepancies at other points of the confidence-scale, though not at zero confidence. In other words, these Os made

⁴E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, 1944.

⁵Four of the Os gave fewer than 400 reports, as indicated in Table I. These Os reported that they were blinking or not looking at the screen when some of the exposures occurred. One, who gave only 371 reports, said that she was having trouble with her vision during the experiment and had to keep her eyes closed part of the time.

many more correct judgments than they thought they were making, though they were essentially right in thinking that their 'pure guesses' were indeed no better than guesses at random. It would seem that the use of the entire scale provides a more sensitive measure of underconfidence than the zero point alone. Furthermore, because of the way in which the scale is defined, every *O* accepted the positive discrepancies as indicating underconfidence in his reports.

Every *O* gave a positive relationship between percentage correct and judgment

TABLE I
PERCENTAGES OF CORRECT JUDGMENTS AT EACH POINT ON THE CONFIDENCE-SCALE
Percentages based on at least 10 judgments are shown in boldface. The number of judgments at each point is shown in parentheses

Confidence	Observers										Total
	1	2	3	4	5	6	7	8	9	10	
100	100 (168)	100 (8)	100 (255)	100 (236)	100 (282)	94 (161)	96 (111)	99 (190)	96 (289)	90 (115)	97.9 (1815)
90	100 (20)	100 (20)	100 (9)	100 (16)	100 (15)	73 (11)	75 (4)	95 (20)	100 (4)	80 (5)	95.2 (124)
80	100 (20)	100 (10)	92 (13)	100 (12)	80 (5)	100 (7)	77 (9)	50 (2)	100 (8)	71 (14)	91.0 (100)
70	90 (20)	100 (17)	— (0)	100 (8)	88 (17)	100 (8)	50 (4)	86 (14)	100 (3)	80 (5)	90.6 (96)
60	100 (13)	100 (19)	100 (2)	75 (4)	— (0)	100 (2)	75 (8)	100 (9)	60 (5)	0 (3)	87.7 (65)
50	100 (24)	100 (37)	94 (17)	88 (17)	92 (13)	91 (22)	77 (18)	59 (17)	56 (25)	42 (12)	82.7 (202)
40	75 (4)	100 (6)	— (0)	59 (17)	100 (1)	29 (7)	55 (9)	55 (11)	0 (3)	50 (12)	55.7 (70)
30	92 (12)	100 (15)	95 (21)	60 (10)	60 (10)	20 (7)	50 (14)	40 (10)	33 (6)	35 (17)	66.4 (110)
20	81 (26)	81 (32)	80 (9)	69 (13)	20 (7)	43 (30)	54 (13)	29 (24)	0 (5)	13 (23)	52.8 (182)
10	54 (41)	34 (82)	65 (23)	13 (67)	19 (16)	11 (72)	17 (59)	15 (53)	0 (1)	11 (64)	23.0 (478)
0	44 (45)	0 (125)	20 (51)	— (0)	3 (34)	1 (73)	2 (146)	0 (50)	0 (44)	0 (130)	4.9 (698)
Total	86.5 (393)	50.1 (371)	86.8 (400)	81.0 (400)	85.3 (400)	55.8 (400)	43.0 (395)	66.0 (400)	79.7 (390)	37.0 (400)	67.2 (3949)

of confidence. This result is in line with former laboratory studies of judgments of confidence.

None of the *Os* was clearly overconfident. Under different experimental conditions, however, overconfidence can easily be obtained.⁶

It should be noted that over- or underconfidence is not simply a function of the *distribution* of the judgments of confidence, though there may be a statistical relationship involved. *Os* 3, 4, and 5 are more underconfident than *Os* 6, 7, 8, and 10, yet their confidence-judgments tend to be much higher. Individual differences in accuracy mask any simple effect of distribution of judgments.

EXPERIMENT II

Procedure. An attempt was made in Experiment II to train *O* in confidence-judgments to see how small the discrepancies between the percentages of correct word-recognitions and the objective meaning of the scale-points could be made.

⁶ Further experiments by Dr. Pauline Austin Adams and the author have produced marked overconfidence.

O was given knowledge of results at the end of each experimental session; *i.e.* he was shown the tally of correct and incorrect reports for each point on the confidence-scale. *E* reminded *O* of his performance at the beginning of the next session. There were five experimental sessions of 1 hr. each. The words were exposed in random order, the same word never being exposed consecutively. There were 165 different words used, each typewritten in capitals and exposed with an opaque projector and

TABLE II

PERCENTAGE OF CORRECT JUDGMENTS AT EACH POINT ON THE CONFIDENCE-SCALE,
FOR ONE *O* DURING FIVE EXPERIMENTAL SESSIONS

Percentages based on at least 10 judgments are shown in italics. The number of judgments at each point is shown in parentheses

Confidence	Experimental session				
	1 (80)	2 (120)	3 (80)	4 (115)	5 (115)
100		100 (12)	100 (11)	100 (6)	
90		100 (2)	89 (9)	100 (8)	
80		100 (15)	100 (8)	90 (10)	100 (1)
70			80 (10)	50 (6)	
60		100 (5)	73 (11)	64 (11)	100 (2)
50		75 (32)	43 (7)	0 (4)	100 (2)
40		100 (3)	33 (3)	14 (7)	100 (3)
30	100 (7)	57 (7)	20 (5)	13 (8)	50 (2)
20	88 (17)	45 (11)	0 (5)	10 (20)	38 (16)
10	61 (28)	55 (11)	20 (5)	6 (31)	23 (62)
0	0 (28)	0 (22)	0 (6)	0 (4)	7 (27)
$\Sigma n_i d_i /T$	35.4	18.2	7.6	6.5	12.2
Total	49	63	61	34	27

shutter. The words differed widely in word-frequency (Thorndike-Lorge L count). Exposure was for about 0.20 sec. *O* guessed verbally after each exposure and then gave his confidence in his judgment, using only the Points 0, 10, . . . , 100%. Exposures were at the rate of about one each 10 sec. Results in terms of percentage correct and deviation from objective meaning are given in Table II. Percentages based upon at least 10 judgments are given in italics. Also shown is the mean absolute discrepancy in percentage for each session. This mean is based upon all discrepancies and is computed by weighting each discrepancy by the number of judgments upon which it is based. The formula used is $\Sigma n_i |d_i|/T$, in which $|d_i|$ is the absolute discrepancy, n_i is the number of judgments upon which $|d_i|$ is based, and T is the total number of judgments made during the given session. That some improvement occurred is indicated by the decrease in the mean discrepancy. The

procedure was tedious, and motivational changes may account for the increase in the mean discrepancy during the last session.

That the mean discrepancy is not solely a function of percentage correct is shown by the last two lines of Table II, in which the lowest discrepancy was achieved along with the next to lowest percentage correct (Session 4). In other words, the conditions which lead to a high percentage correct are not identical with the conditions leading to accurate judgments of confidence. Nevertheless, we cannot be sure, even for this one *O*, that whatever external cues he was responding to in making more accurate judgments of confidence are not the same cues he was responding to in reporting the words seen. The reason for the latter statement is that there were slight (and unknown) variations in shutter-speed from session to session which may account for most of the variation in percentage correct, as well as for at least some of the variation in discrepancy-score.¹ Replication is underway with better control of experimental conditions.

Further experimentation now underway involves the use of the scale in the study of variables related to overconfidence and underconfidence and in the training of *O*s to make more realistic confidence-judgments, including the study of transfer of training.

¹ Later experiments by Dr. Pauline Austin Adams and the author have shown that improvement in judgments of confidence cannot be accounted for by improvement in performance (percentage correct) when external cues are held constant. These experiments, however, deal with very different kinds of tasks.

SUBLIMINAL STIMULATION IN PROBLEM-SOLVING

By PAUL A. KOLERS, V.A. Hospital, West Haven, Connecticut

The phenomenon of metacontrast occurs as follows: If two contours are presented serially and tachistoscopically at appropriate time-values, the first is not reported. If, however, the first is presented alone, it always is reported. The phenomenon thus consists of an apparent inhibition of the first contour by the second.¹ Taking advantage of this phenomenon, several investigators have explored the influence of a subliminal first (*A*) figure upon the perception of a supraliminal second (*B*) figure.² For example, Smith and Henriksson presented fan lines as the *A*-figure and a square as the *B*-figure. Their *Os* did not report seeing the fan lines but tended to describe the square as trapezoidal. Similarly, Klein *et al.* presented representations of genitalia as *A*-figure and sexually ambiguous forms as *B*-figure, finding that the descriptions of the forms (*B*-figure) tended to be congruent with the sexuality of the unreported *A*-figure.

It would not be difficult to point to a relationship between these studies and recent investigations of 'perceptual defense' and related phenomena. One characteristic of such studies is the restricted range of stimulus-materials that have been used; generally, these studies have required only an immediate response identifying or attributing characteristics to a stimulus. Apparently, no effort has been made to determine whether subliminal stimuli may affect more complex cognitive processes, such as those presumably involved in the perception of relations and in problem-solving. The present paper reports two such experiments.

METHOD

Presentation. The intensity and duration of the stimulus-figures, and the delay between them, were arranged at values that combined a good inhibitory effect with maximal 'registration' of the *A*-figure. Problem-cards were presented as *B*-figure in the tachistoscope, and information relevant to solving the problems as *A*-figure. The *A*-figures were presented on the same part of the mirror of the tachistoscope that was subsequently occupied by the horizontal parallel lines in Fig. 1. This method is deemed particularly suitable for the study of some aspects of problem-solving since it combines manipulation of *S*'s visual field without his awareness with the control associated with tachistoscopic projection.

* Received for publication March 4, 1957. Abstracted from a doctoral dissertation submitted to the faculty of the Graduate School of Arts and Science at New York University. The writer is grateful to Dr. D. P. Spence, to Dr. Richard Christie, and to Dr. M. D. S. Braine for advice.

¹ For historical background see M. Alpern, Metacontrast: Historical Introduction, *Amer. J. optom. Arch. Amer. Acad. Optom.*, 29, 1952, 631-646; H. H. Toch, The perceptual elaboration of a stroboscopic presentation, this JOURNAL, 69, 1956, 345-358. See also Heinz Werner, Studies on contour: I. Qualitative analyses, this JOURNAL, 47, 1935, 40-64; P. G. Cheatham, Visual perceptual latency as a function of stimulus brightness and contour shape, *J. exp. Psychol.*, 43, 1952, 364-380.

² G. J. W. Smith and Maj Henriksson, The effect on an established percept of a perceptual process beyond awareness, *Acta Psychol.*, 11, 1955, 346-355; G. S. Klein, D. P. Spence, R. R. Holt, and Susanna Gourevitch, Preconscious influences upon conscious cognitive behavior, *Amer. Psychol.*, 10, 1955, 80 (Abst.); U. Kragh, *The Actual-Genetic Model of Perception-Personality*, 1955, 1-398.

Stimulus-objects. Two different classes of stimulus-objects were constructed to test problem-solving at two different 'levels' of difficulty. In the first of these (simple abstraction), *S* was shown a series of cards on which were drawn three rows of geometrical figures, and instructed to name one figure that appeared in all three rows. Two figures appeared in all three rows, and one of these was flashed as *A*-figure to half the *Ss*. Fig. 1 (left half) shows one of the problem-cards used in Experiment I.

The second class of stimulus-cards presumably test abstraction and generalization. *S* was shown cards on which there were two rows of figures, three in the top and

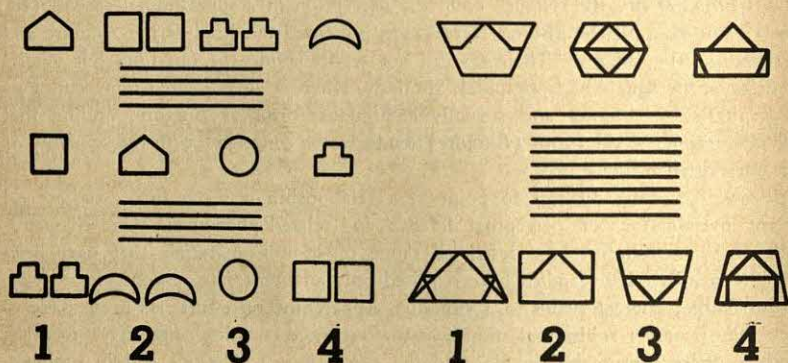


FIG. 1. SAMPLE PROBLEM-CARDS USED IN EXPERIMENTS I (LEFT) AND II (RIGHT). In Experiment I, squares were exposed to half the *Ss* as *A*-figure and appeared in the area indicated by the horizontal lines. In Experiment II, the form common to the figures of the top row and to No. 1 in the bottom row was given as *A*-figure to half the *Ss* and appeared in the area occupied by the horizontal lines.

four in the bottom. The three figures in the top row had in common a simple geometrical form; one of the figures in the bottom row also had this element in it. *S* was instructed to name the figure in the bottom row that contained the element present in the three figures of the top row. This element was flashed as *A*-figure to half the *Ss*. Fig. 1 (right half) shows one of the problem-cards used in Experiment II.

Apparatus. A 3-field tachistoscope was built for these experiments. The viewing box was of a modified Dodge type, the light-sources for which were six cold-cathode tubes coated with magnesium tungstate phosphor. The mirrors, coated with titanium oxide, have a 60:40 transmission-reflection ratio $\pm 5\%$.³ Field 1 of the viewing box was used for light-adaptation; stimulus-materials were placed in Fields 2 and 3 (*A*-figure in 2, *B*-figure in 3). The visible fields measured 9×9 in. and were at an optical distance of 34 in.

The apparatus could be triggered manually or made to cycle automatically. For manual operation, Field 1 was on, then Fields 2 and 3 were lighted sequentially for predetermined durations, then Field 1 was lighted again. The basic cycle was thus

³ The light tubes can be obtained from Aristo Grid Lamp Products, Inc., of Port Washington, New York; the mirrors from Chemoptics, Inc. of New York City.

Fields 1-2-3-1. For automatic operation, this basic cycle was repeated until stopped manually. There was no discriminable pause between the extinction of one field and the lighting of the next. The circuit operated off high speed, high voltage *DPDT* relays. Exposure-durations used for an *S*, once determined, were not changed during the course of the experiment, and the slight noise the relays made could therefore provide no information about the nature of the stimulus-card. A separate unit independently controlled the brightness of each phosphor-coated tube, which insured equivalent illumination of the three fields despite variations in reflection-transmission characteristics of the transparent mirrors. This illumination, measured at the viewing hood with a Macbeth Illuminometer, was 1.4 ft.-C.⁴

EXPERIMENT I

Subjects. The *Ss* of Experiments I and II were 40 undergraduate students at New York University, 20 men and 20 women. They were assigned in prearranged random order, 5 *Ss* to each cell of a $2 \times 2 \times 2$ table. All *Ss* reported normal vision or vision corrected to 20:20.

Procedure. The variables of the $2 \times 2 \times 2$ table were Sex, Condition, and Order. Sex was chosen because of recent reports of sex-differences in perceptual functioning.⁵ In one condition, *S* was given an answer subliminally; in the other condition a blank field was presented as *A*-figure. Order refers to experimental sequence: half the *Ss* took part in Experiment II before Experiment I, and the other half in Experiment I before Experiment II.

The experimental room was dimly illuminated by a 25-w. lamp in a goose-neck fixture. Efforts were made to create an informal atmosphere because pilot-studies seemed to indicate that this was an important variable. The *Ss* were shown a number of geometric forms to acquaint them with tachistoscopic exposure. Exposure-durations were: an initial 30 sec. of light-adaptation followed by 0.03 sec. of *A*-figure, 0.30 sec. of *B*-figure and 2.0 sec. of blank field. Reports were given during the 2-sec. blank-field exposure between cycles.

One of the 6 problem-cards used appears as Fig. 1. The critical part of the instructions given to *S* was as follows:

You will see cards with three rows of geometrical figures on them, four figures to the row. The bottom row of figures is numbered 1,2,3,4. Find in the bottom row, and identify by its number, one figure that appears in the other two rows as well. Call out this number after each exposure of the card.

S was given two problems with blank *A*-figures as practice. During the first five exposures of both the practice-cards and the test-cards, the bottom row of figures was hidden by a plain gray card. This was done to insure that *S* saw the problem before he attempted to answer it. Beginning with the sixth exposure, the gray card was not present, and *S* called out numbers that represented his 'best guess,' for each exposure, of the right answer. Exposures were discontinued when either of two criteria was met: (1) *S* called out the same answer for 10 consecutive exposures; or (2) failure to give 10 consecutive identical answers within 25 exposures of the problem-card.

⁴ A full description of an improved version of the device is being prepared for publication.

⁵ For example, H. A. Witkin, H. B. Lewis, M. Hertzman, K. Machover, P. B. Meissner, and S. Wapner, *Personality through Perception*, 1954, 1-571.

There are two objectively correct answers to each card. One of these answers was flashed as *A*-figure to half the *Ss* (experimental group); a blank field was flashed as *A*-figure to the other half (control group). Interest is in the number of times the groups chose that one of the two objectively correct answers which was congruent with the *A*-figure presented to the experimental group. For statistical analysis, these congruent answers were arbitrarily designated correct. To demonstrate that the *A*-figures affected *S*'s choice, it must be shown that the experimental group chose

TABLE I
SUMMARY OF MEANS IN EXPERIMENT I

<i>Ss</i>	Order I		Order II	
	experimental	control	experimental	control
Men	1.886	2.020	2.144	1.616
Women	2.252	1.804	1.854	1.754

correctly more frequently than did the control group. Scores could range from 0-6. If, for example, *S* gave objectively correct answers on all six cards, but not the ones that *E* flashed to half the *Ss*, his score was 0. The higher *S*'s score, the more frequently did his responses resemble answers being flashed to half the group subliminally.

Results. Because of the small counts involved, the scores were transformed to $(X + 0.5)^{1/2}$.⁶ The summary of means of the transformed data will be found in Table I. Bartlett's test of these data indicated that homogeneity of variance existed. The analysis of variance of the data yielded *F*s of 7.148 and 8.217 for condition and for the second-order interaction, respectively (for 1 and 32 *df.*, $p < 0.025$). The interpretation of the condition-variance must be cautious because of the significant interaction and the reversal of trend that occurs in the first comparison (men) shown in Table I, the causes of which are unknown.

EXPERIMENT II

The same *Ss*, exposure-durations, and general procedure were used in Experiment II as in Experiment I. Control and experimental *Ss* of the first experiment reversed roles in the second. One of the six problem-cards used appears as Fig. 1 (right half). The critical part of the instructions given to the *Ss* was as follows:

You will see cards with two rows of geometrical figures on them, three figures in the top row and four numbered figures in the bottom row. The three figures in the top row have an element in common—some part occurs in all three. One of the figures in the bottom row also has this element in it. Your task is to identify by number the one figure in the bottom row that has in it the element present in all three figures of the top row. Call out this number after each exposure of the card.

There is one objectively correct answer to each card. Part of this answer was flashed as *A*-figure to half of the *Ss* (experimental group); a blank field was flashed as *A*-figure to the other half (control group). The score of each *S* was the number of problems he actually solved; and scores could thus range from 0-6.

Results. Again the scores were transformed to $(X + 0.5)^{1/2}$. Table II shows the

⁶ A. L. Edwards, *Experimental Design in Psychological Research*, 1950, 200.

means and variances of the transformed data. Bartlett's test of these data indicated homogeneity of variance existed. The analysis of variance of the data yielded F s of 15.07 and 7.39 for order and condition, respectively (for 1 and 32 df , $p < 0.025$). That is, S s who received the relevant contour as A -figure solved more problems than did the control S s; and S s who took part in Experiment II after Experiment I solved more problems than S s in the reverse sequence.

An equally important analysis of these data would concern rate of solution—the number of exposures required. Such an analysis could not be carried out reliably for two reasons: (1) the problems used were not equally difficult; (2) no single problem was solved by all S s. A problem-by-problem analysis based on variable n s seems to suggest, however, that the effect of the subliminal stimulus is equivocal: while

TABLE II
SUMMARY OF MEANS IN EXPERIMENT II

Ss	Order I		Order II	
	experimental	control	experimental	control
Men	2.054	1.352	2.338	2.066
Women	2.016	1.836	2.206	2.202

all S s solve more problems when given this information, the rate of solving appears to be slowed down in men and speeded up in women when they are compared with their respective control groups. Similar results were found in pilot-studies.

Some years ago, Maier showed that S s who had not been able to solve the pendulum-problem after a 10-min. inspection solved it in 42 sec. (average) after Maier brushed against one of the strings and gave the hint that it needed to be made to swing.⁷ Of particular interest is the fact that none of the S s given the hint reported awareness of it, either at the time of solution or later. The import of the present experiments may, therefore, be regarded as similar to that of Maier's; however, the introduction into the tachistoscope of both the problem and the cue to its solution constitutes a methodological advance over earlier work, for the tachistoscope permits carefully controlled manipulation of both problem- and cue-variables. With the demonstration of the basic effect, the way is opened to parametric studies of some of the variables involved in complex cognitive processes. One of the most obvious approaches is the introduction of irrelevant and erroneous A -figures at different stages in the development of the solution.

SUMMARY

The problem under consideration was the influence of subliminal stimulation in problem-solving. Two classes of problem were presented tachistoscopically, and, for each class, half the S s were given cues to solution subliminally, using the phenomenon of metacontrast. The results indicate that the subliminal stimuli influenced the kind and number of solutions given.

⁷ N. R. F. Maier, The solution of a problem and its appearance in consciousness, *J. comp. Psychol.*, 12, 1931, 181-194.

VERBAL AND INSTRUMENTAL RESPONSE-HIERARCHIES AND THEIR RELATIONSHIP TO PROBLEM-SOLVING

By ARTHUR W. STAATS, Arizona State College at Tempe

There have been a number of recent attempts to account for human problem-solving in *S-R* terms.¹ Osgood, for example, has suggested that *S* learns hierarchies of perceptual responses to objects which mediate hierarchies of instrumental responses, with the result that problem-solving behavior can be predicted from verbal responses.² Thus, the latency with which *S* states how an object might be used would indicate the latency of the actual use of the object in a problem-situation. Osgood did not discuss the possibility that there might be a difference between the verbal and the instrumental hierarchies, although it seems quite possible that the two classes of responses might be acquired independently. A study by Saugstad has indicated that there is at least a gross relationship between verbal and instrumental responses to objects in a problem-solving situation, but he was not interested in relative verbal response-strengths, e.g. latency.³

In the present study, possible verbal uses of a screwdriver were given by *S* before and after attempts to solve a problem which required the use of that implement as a weight. The verbal response measures were related to the latency of the response to the screwdriver as a weight in the problem-situation. Other relationships were also studied: (a) between two different indices of verbal response-strength;⁴ (b) between verbal fluency and problem-solving ability;⁵ (c) between verbal fluency and anxiety;⁶ (d) between sex and problem-solving ability;⁷ and (e) between performance on the Abstract Reasoning Test of the Differential Aptitude Tests and both problem-solving and verbal fluency.

PROCEDURE

Subjects. Sixty-one *Ss* were used, 40 women and 21 men. The *Ss* were under-

* Received for publication March 18, 1957. This study is based on a doctoral dissertation submitted to the University of California at Los Angeles in 1955. The writer is grateful to Professor Irving Maltzman and Carolyn K. Staats for assistance on this work.

¹ C. N. Cofer, The role of language in human problem solving, Conference on Human Problem Solving Behavior, New York University, 1954 (unpublished); Irving Maltzman, Thinking: From a behavioristic point of view, *Psychol. Rev.*, 62, 1955, 275-286; C. E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 603-637.

² Osgood, 631-632.

³ Per Saugstad, Problem solving as dependent upon availability of functions, Unpublished Doctoral dissertation, University of Chicago, 1952, 28-35.

⁴ A. A. Campbell (The interrelations of two measures of conditioning in man, *J. exp. Psychol.*, 22, 1938, 225-243) has found correlations between non-verbal measures.

⁵ Harold Guetzkow (An analysis of the operation of set in problem solving behavior, *J. gen. Psychol.*, 45, 1951, 219-244) has also found such a relation.

⁶ J. A. Taylor (Drive theory and manifest anxiety, *Psychol. Bull.*, 53, 1956, 303-320) has summarized evidence which suggests that the relationship should be positive.

⁷ Guetzkow (*op. cit.*, 219 f.) and N. R. F. Maier (An aspect of human reasoning, *Brit. J. Psychol.*, 24, 1933, 144-155) have reported masculine superiority. If verbal and instrumental functions are related, the same superiority should be found at the verbal level.

graduate students in general psychology at the University of California at Los Angeles.

Apparatus. Two strings were so hung from the ceiling in diagonal corners of a room that *S* could not reach one string while holding the other. The problem-solving objects, arranged close together on a table, were, from left to right: a 1×4 -in. sheet of balsa wood, a 1-in. screw, a $7\frac{1}{2}$ -in. screwdriver, a 2×3 -in. piece of paper, and a 3-in. pencil.

Method. *S* was first given the Abstract Reasoning Test, then verbal hierarchies of functional responses to a pencil, a screw, and a piece of paper were obtained—*S* simply being asked to list all of the different ways that each object might be used (with 5 min. given for each object). *S* was then given the Manifest Anxiety Scale and dismissed. During the following week *S* participated in the problem-solving phase of the study. Before the problem was given, *S* listed uses for the other problem-solving objects—screwdriver, balsa wood, and piece of string. Each *S* was allowed 10 min. in which to solve the problem, *i.e.* to get one string in each hand. To accomplish this, the screwdriver had to be tied to one string and swung in pendular fashion. Then *S* was again shown the screwdriver and given 5 min. to list its uses, following which he was asked to clarify any uses given which were ambiguous. *S* was cautioned not to discuss the experiment with other members of his classes.

Measures. The measure of problem-solving ability employed was time to solution of the problem. Six of the 61 *Ss* did not solve the problem in the allotted time, *i.e.* 10 min. The distribution of times was skewed, but that for log time appeared essentially normal.

A 'weight' response to the screwdriver was defined as a verbal response which depended only upon the weight-property of the screwdriver, not its length, rigidity, sharpness, and so forth, *e.g.* fishing sinker, ballast, paperweight. When the response-measures were obtained subsequent to problem-participation, a response was not counted as a weight-response if it referred specifically to the use of the screwdriver to bring two strings together.

The rank of the first weight-response in the verbal response-hierarchies was considered the latency of the response. If *S* did not give a weight-response, his rank was the number of responses *plus* 1. The distribution of this measure, to be called Latency Weight-Response, was skewed. Frequency of the weight-type of response was the other measure of response-strength (Frequency-Weight-Responses). Its distribution also was skewed.

Another category of response, 'hammer'-response, was defined to include responses which specified that the screwdriver was held by the shaft and so swung that the handle struck and applied force to some other object. Again two indices of the strength of the response-tendencies, latency and frequency of the hammer-responses, were tabulated. Both distributions were somewhat skewed.

The total number of pre-problem verbal responses made to each of the problem-solving objects was tabulated. There was a measure for each of the objects as follows: Total Screwdriver-Responses (pre- and post-problem), Total String-Responses, Total Screw-Responses, Total Wood-Responses, Total Paper-Responses, and Total Pencil-Responses. The last four were summed to yield another measure, Total Responses Other-Objects. These measures were distributed in a somewhat skewed manner, as were the raw scores of the anxiety- and reasoning-tests which were used.

RESULTS

Weight-response and problem-solving. Only 7 of the 61 Ss gave a verbal weight-response in their pre-problem hierarchies, yet all but 6 Ss solved the problem. Thus, contrary to Osgood's expectation and Saugstad's results, in this situation it did not appear feasible to predict problem-solving performance from verbal responses to the problem-object. This lack of relationship may indicate that the verbal responses do not closely parallel the instrumental responses which are learned.

Time required for the solution of the problem was correlated with the two weight-response measures obtained subsequent to problem-participation. Since all three of these measures were correlated with Total Screwdriver Responses, the influence of this variable was controlled by partial correlation. The correlation between log time and latency of the Weight-Response was 0.25 ($p = 0.053$, two-tailed). The correlation of log time and frequency of the Weight-Responses was -0.27 ($p = 0.05$, two-tailed). Although the distribution of Frequency Weight-Responses was skewed, the assumption necessary for the use of the product-moment r appeared to be met, *i.e.* when the mean times of Ss with 1, 2, 3, . . . n weight-responses were calculated, the relationship appeared to be linear. The assumption of linearity appeared to have been met also in all other uses of the statistic.

Thus, although verbal weight-responses were not frequently given prior to problem-participation, the Ss did frequently respond to the screwdriver as a weight subsequent to the problem. This result did demonstrate a relationship between the instrumental response-hierarchy and the subsequent verbal response-hierarchy, *i.e.* between a reinforced instrumental use of the screwdriver as a weight and subsequently verbalized weight-responses. This result has significance for recent studies of verbal behavior which have demonstrated that reinforcement can affect classes of verbal responses.⁸ Furthermore, these verbal responses were emitted in strength proportional to the time it took the Ss to solve the problem, a relationship not explained by the effect of problem-participation upon the subsequent verbal weight-response, since only one reinforcement was involved whether the problem was solved quickly or slowly. The difference in the strength of verbal weight-responses which the Ss later demonstrated might have been learned prior to the problem. Problem-participation, with the reinforcement of the instrumental use of the screwdriver, might have brought the class of verbal weight-responses into a more dominant position for all the Ss who solved the problem. This interpretation suggests that, if the verbal weight-responses could have been elicited prior to problem-participation, the strength of the responses would have been positively related to time. Notwithstanding the interpretation, however, the results show only a relationship between problem-solving participation and subsequent verbal responses.

Latency and frequency. Latency and frequency of the Weight-Responses were correlated -0.67 ($p = 0.001$, two-tailed). Again the influence of Total Screwdriver-Responses was controlled by partial correlation. The partial correlation of latency and frequency of the Hammer-Responses was -0.76 ($p = 0.001$). The results corroborate the expectation that the different indices of verbal response-strength would be correlated.

⁸ B. D. Cohen, H. I. Kalish, J. R. Thurston, and E. Cohen, Experimental manipulation of verbal behavior, *J. exp. Psychol.*, 47, 1954, 106-110; W. S. Verplanck, The control of the content of conversation: Reinforcement of statements of opinion, *J. abnorm. soc. Psychol.*, 51, 1955, 668-676.

Fluency and problem-solving. The correlations of the fluency-measures and log time are shown in Table I. A greater number of responses was associated with more rapid problem-solution.

Fluency and anxiety. The Manifest Anxiety scores were correlated with the total number of responses given to each of the objects involved in the problem-situation, as shown in Table I. In each case, the correlation was positive, but in only two cases were they statistically reliable.

Sex-differences. As expected, men solved the problem more quickly than women. The Mann-Whitney *U*-test of this difference yielded a *Z* of 2.95 ($p = 0.003$). Men

TABLE I
CORRELATIONS OF FLUENCY-MEASURES WITH LOG TIME AND MANIFEST ANXIETY

Total responses	Anxiety	<i>p</i>	Log Time	<i>p</i>
Screwdriver (pre-problem)	0.17	—	-0.53	0.001
Screwdriver (post-problem)	—	—	-0.52	0.001
String	0.30	0.01	-0.47	0.001
Screw	0.21	0.105	—	—
Wood	0.05	—	—	—
Paper	0.26	0.05	—	—
Pencil	0.14	—	—	—
Other objects (screw, wood, paper, pencil)	—	—	-0.28	0.05

also gave more responses to the screwdriver.⁹ On the Mann-Whitney *U*-test, the *Z* was 2.79 ($p = 0.005$). Men gave their first hammer-response lower in their verbal response-hierarchies (pre-problem). The *Z* was 4.70 ($p < 0.0001$). Women gave more hammer-responses. To analyze the results, men and women were classified into two categories, those who gave no hammer-responses, and those who gave one or more hammer-responses. Twelve men were in the first category, 8 in the second. For women the tabulation was 9 and 32. The X^2 was 8.52 ($p < 0.005$). Thus, both indices indicated that women had learned a stronger hammer-response to the screwdriver. The corresponding comparisons on the strength of the weight-response yielded (a) a *Z* of 0.76 ($p = 0.45$) and (b) a X^2 of 0.08 ($p = 0.80$). No sex-difference was shown.

These results demonstrate that the responses learned to the screwdriver differ for the sexes in quantity and quality. Women appeared to have learned a stronger hammer-response to the screwdriver,¹⁰ even while giving fewer total responses. Thus, it may be inferred that men and women *Ss* have different experience with respect to the screwdriver, and that this difference in experience may be reflected in their differential ability to solve the 'two-string' problem.

Other results. (a) Men and women did not differ significantly on the Abstract Reasoning Test; the Mann-Whitney *U*-test yielded a *Z* of 0.52 ($p = 0.60$). (b) Scores on the Abstract Reasoning Test and log time correlated -0.33 ($p < 0.01$). (c) Latency and frequency of Weight-Response correlated -0.02 ($p = 0.88$) and -0.05 ($p = 0.70$) with abstract reasoning, respectively. Thus, more intelligent *Ss*

⁹ To save space, the sex-differences in the number of responses made to the other problem-solving objects are not presented here. None of the differences was significant, although in each case men gave more responses.

¹⁰ To save space, the inconclusive and cumbersome analyses of the relationships between strength of hammer-response and problem-solving are not presented. Some evidence, however, suggested a negative relationship.

did not give stronger post-problem weight-responses to the screwdriver than did less intelligent ones. (d) The correlation of abstract reasoning with Total Screwdriver-Responses was 0.05 ($p = 0.70$). Thus, fluency, an asset in problem-solving, was not accounted for by individual differences in intelligence. (e) Manifest Anxiety was not significantly correlated ($-0.11, p = 0.41$) with log time.

SUMMARY

Verbal responses to objects used in problem-solving were obtained before and after problem-participation. (a) The results indicated no relationship between pre-problem verbal responses and problem-solution, but (b) there was a relationship between rapidity of problem-solution and the strength of verbal responses subsequently given. (c) The two indices of strength of verbal response which were used were significantly correlated. (d) Fluency of verbal responding was positively related to problem-solving. (e) Positive relationships among anxiety-level and fluency-measures appeared, but four of the six correlations were not significant. (f) High scores in Abstract Reasoning went with rapid problem-solution, but not with fluency. (g) Men exhibited superior ability in problem-solving but evidence indicated that the difference may have been due to differential experience with the objects used in the problems. (h) The sexes did not differ in abstract reasoning.

ASSOCIATION OF NONSENSE-FIGURES AS A FUNCTION OF FITTINGNESS AND INTENTION TO LEARN

By JACQUES KASWAN, Yale University

The point of departure for the present study is the assumption that perceptual organization facilitates retention.¹ The assumption is supported by Köhler's finding of better recall for similar as compared with dissimilar items and for 'fitting' as compared with 'non-fitting' items in paired-associates experiments.² Köhler also suggests that association is spontaneous when the structure of the learning materials is favorable to organization, while poorly structured materials require a "particular attitude" of "active organization" if associations are to be formed.³ The validity of this formulation was examined in the present experiment, which deals with the retention of fitting (*F*) and non-fitting (*NF*) pairs of nonsense-figures (of comparable intra-pair similarity) in intentional and incidental learning. A finding of superior recall for *F* over *NF* pairs under both conditions would be evidence for some generality and stability of fittingness as a variable of organization. Köhler's analysis further suggests that under the incidental condition, which minimizes "active organization," the effect of fittingness should be relatively greater than under the intentional condition.

Materials. There were 16 pairs of nonsense-figures drawn in black ink on white 4 × 6-in. cards, each figure measuring approximately 3 × 4 cm., with a space of 0.5 cm. between the members of each pair. The adjacent sides of the two members of each pair of *F*-figures were designed to match exactly—if brought together they would form a continuous whole. Phenomenologically, these figures appeared to be fitting. For the purposes of this study, then, fittingness is primarily a function of "good continuation" as described by Wertheimer.⁴ There were four outline and four solid pairs arranged to form eight *F* and eight *NF* pairs. The 16 pairs of figures were divided into two series, Series A (pairs 1-8) and Series B (pairs 9-16), each consisting of four *F* and four *NF* pairs (Fig. 1). The two series are made up of identical items—the left-hand (stimulus) items are the same and the right-hand (response) items are the same. The only difference between them is in the pairing of stimulus- and response-items. Items which are *F* in Series A are *NF* in Series B, and those which are *NF* in Series A are *F* in Series B.

Procedure. The *Ss* of the incidental group were told that the experiment was concerned with extra sensory perception (*ESP*). They were asked to concentrate on the figures shown for the benefit of a 'receiver' seated in a corner of the room.

* Received for publication September 28, 1956. This report is based on a thesis submitted to the Graduate Faculty of Political and Social Science of the New School for Social Research in partial fulfillment of the requirements for the M.A. degree. The writer is indebted to Professor Hans Wallach, who directed the research, and to Professors Mary Henle and Irving Rock for valuable suggestions and criticisms. Professors John Flynn and Robert Abelson of Yale helped resolve statistical problems.

¹ Wolfgang Köhler, *Gestalt Psychology*, 1947, 270-271.

² Köhler, On the nature of association, *Proc. Amer. Phil. Soc.*, 84, 1941, 489-502.

³ Köhler, *Gestalt Psychology*, 271.

⁴ Max Wertheimer, Laws of organization in perceptual forms, in W. D. Ellis (ed.), *A Sourcebook of Gestalt Psychology*, 1950, 71-88.

Half the Ss were shown the cards of Series A, and the rest those of Series B. Each card was exposed for 3 sec., with an interval of 2 sec. between successive cards. There were two presentations of the series without pause, the order of presentation

SERIES A SERIES B SIMILARITY SERIES

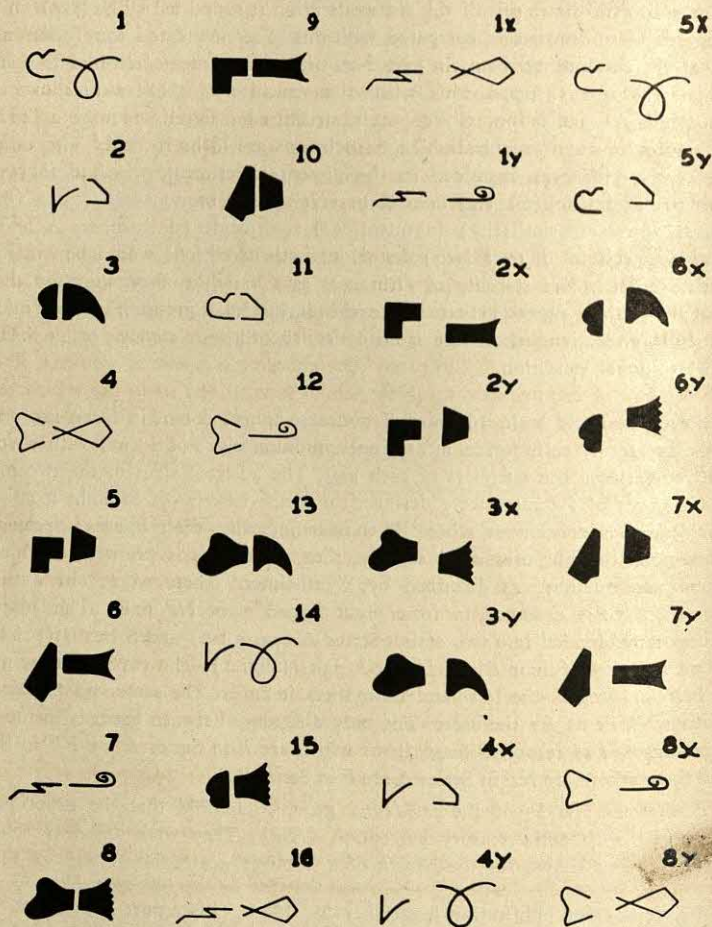


FIG. 1

being varied from S to S according to an 8×8 arrangement. Different orders were used for training and testing to minimize serial bias. There was a 5-min. interval between training and testing during which E conversed with S, warning him not to think of the cards for fear of disturbing the 'receiver,' who was supposedly writing an introspective report. Retention was tested by presenting the left-hand figure of each pair previously shown and asking for a reproduction of its right-hand figure

on separate 4×6 in. slips of paper. If *S* did not begin to draw within 10 sec. after the presentation of each stimulus-card, the next card was given. A maximal period of 25 sec. was allowed for the drawing of each figure. For the *intentional* (*INT*) group, the procedure was the same as for the *ESP* group, with the exception that *Ss* were not given the *ESP* cover-story but were told merely to memorize the figures in preparation for a recall test.

The intra-pair similarity of the materials was estimated as follows: All 16 pairs of figures were drawn on another set of cards. The new cards were different only in that the distance between the members of each pair was increased to minimize fittingness (Fig. 1). Upon conclusion of the recall test, all *Ss* were shown simultaneously the *F* and *NF* pairs with the same left-hand figure and were asked which two figures in each pair resembled each other more closely. They were asked to give reasons for each decision, but were not encouraged to decide in favor of either pair if they felt that there was no basis for judgment.

Subjects. A total of 69 *Ss* were tested. The results of five in the incidental group were discarded before evaluation (when upon routine inquiry they expressed suspicion about the purpose of the experiment) leaving 32 in each group. There were 35 men and 20 women, ranging in age from 16 to 49 yr., with a mean of 24.3. The *Ss* were recruited in a variety of places (neighboring schools, a women's residence club, a place of employment, the New School lounge, and so forth) which resulted in a wide range of backgrounds and minimal sophistication in psychology. In view of the nature of the experiment, only one condition was studied with *Ss* drawn from each source.

Scoring. Responses were scored by comparing them with the actual designs. Two different methods of evaluation were used to check the possible bias which scoring by one individual might produce. (1) The author compared responses with the correct designs to determine whether each reproduction conformed to the true figure in important details. In cases of doubt the response was scored incorrect. (2) The second scorer was an intelligent layman. For him the pairs were broken up and the 16 individual figures were mounted on separate cards. The scorer was instructed to decide whether an *S*'s recall-drawing resembled any of the 16 figures. The responses were supposed to represent only right-hand figures and the presence of the left-hand figures therefore decreased the influence of any possible bias, especially since the scorer was not acquainted with the purpose of the experiment. Any response which was judged to resemble two or more figures equally well was scored incorrect. These two methods of evaluation gave 93.8% agreement. The results are presented as determined by the first method, which was selected by the toss of a coin.

Results. The mean difference between *F* and *NF* pairs recalled was 1.04 for the *ESP* group and 0.93 for the *INT* group. Both differences favored the *F* pairs and were significant beyond the 0.1% level of confidence (*t*-test, one tail). As might be expected, total recall was significantly higher under the *INT* than under the *ESP* condition. A comparison of recall for each pairing for both conditions yields a rank correlation of 0.52, indicating fair stability of the effect under the different attitudes induced. A Kolmogorov-Smirnov analysis of the difference between the two groups with respect to proportion of *F* items in the total recall ($F/F + NF$)

gave significance beyond the 10% level.⁵ While in the expected direction, this result does not unequivocally support the expectation that fittingness is a more effective organizing process under the incidental condition. It is to be noted, however, that since the *Ss* were urged to concentrate hard on the designs during this condition, the results may reflect different effects of 'active' organization, rather than a difference between spontaneous organization and patterning as a result of effort.

The results of similarity judgments show that left- and right-hand figures were no more similar to each other in *F* than in *NF* pairings. There were 203 judgments in favor of *F* pairs and 209 in favor of *NF* pairs, while in 100 cases *Ss* declined to choose. This method, though inadequate for objective definition, is widely used to measure similarity,⁶ and provides sufficient evidence for this study that our fittingness variable is independent of similarity, as judged.

Conclusion. The results show that fittingness, produced primarily by the factor of good continuation, appears to induce a relationship between figures which persists in time—as evidenced in facilitation of recall—regardless of intention to learn. While the results are equivocal regarding the increased facilitating effect of fittingness under the incidental condition, they are interpreted to support Köhler's postulation of association as an after-effect of perceptual organization.

⁵ Leo Goodman, Kolmogorov-Smirnov tests for psychological research, *Psychol. Bull.*, 51, 1954, 160-168.

⁶ C. I. Hovland, Human learning and retention, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 663.

THE EFFECT OF DIFFERENT AMOUNTS OF ALTERNATING PARTIAL REINFORCEMENT ON RESISTANCE TO EXTINCTION

By E. J. CAPALDI, The University of Texas

In a recent experiment using rats as Ss, a greater amount of training under an alternating schedule of partial reinforcement led to significantly faster extinction than a lesser amount of such training; while an increased amount of training under a 50% random schedule of reinforcement resulted in about the same degree of resistance to extinction as the lesser amount.¹ This finding suggests that rate of extinction is a function of Ss' ability to discriminate between trials (training and extinction) and the discrimination may be based on the appreciation of sequences of patterns of events in training. That increased amounts of training under an alternating schedule of partial reinforcement would affect discrimination was inferred from several recent experiments.² These studies demonstrated that a *patterning effect* appears when Ss are trained with alternating partial reinforcement. In the case of an instrumental running response, the patterning effect consisted in the latencies of reinforced trials being below the latencies of nonreinforced trials. Elsewhere the occurrence of the patterning effect has been taken as behavioral evidence for increased learning of the pattern of reinforcement itself.³ Granting better learning of the pattern of reinforcement when the patterning effect appears, then it follows that deviations from the pattern can be better discriminated. The experiment to be reported was designed to examine the relation between resistance to extinction and amount of alternating partial reinforcement in human Ss.

Apparatus. The apparatus can best be described as a modification of the Wisconsin 'monkey-apparatus.' It consisted of a partition which obscured S's view of E, a movable tray which could be pushed toward or drawn away from S, with a single well in the tray covered by a lid. A guillotine-door placed in the center of the partition, when raised, permitted the tray to be freely drawn; when closed, it permitted E to bait the well without S's knowledge. The well was baited with a piece of red velvet cloth. The contact of the cloth with the well was noiseless. The entire apparatus was painted gray.

Subjects. The Ss were 45 men and women, volunteers from various classes in introductory psychology, who were assigned to one of three groups of 15 Ss each in ABCCBA order.

Procedure. S was seated before the apparatus and given the following instructions:

When this door is raised [E points to and raises the door], this tray [E points to the tray and jiggles it back and forth] can be pushed toward you. This lid conceals a little well which you can see now that I have lifted the lid. You are to tell me if you think this [E shows S the piece of red velvet cloth] is under the lid when the tray is pushed toward you. Give the best estimate you can. If your estimate is that the cloth is under the lid, say "in"; if your estimate is that the cloth is not under the lid, say "out." After having given your estimate, lift the lid to determine if

* Received for publication February 28, 1957.

¹ E. J. Capaldi, The effect of different amounts of training on the resistance to extinction of different patterns of partially reinforced responses. *J. comp. physiol. Psychol.*, in press.

² E. D. Longenecker, John Krauskopf, and M. E. Bitterman, Extinction following alternating and random partial reinforcement, this JOURNAL, 65, 1952, 580-587; D. W. Tyler, E. C. Wortz, and M. E. Bitterman, The effect of random and alternating partial reinforcement on resistance to extinction in the rat, *ibid.*, 66, 1953, 57-65.

³ Capaldi, *op. cit.*

you were correct. Then replace the lid. Remember, I want the best estimate you can make under the circumstances. Do you understand?

Acquisition. Three groups of Ss were used. Group 1 found the cloth under the lid only on Trial 2, never before or after. Group 2 found the cloth under the lid on Trials 2, 4, and 6, but not on any other trial. Group 3 found the cloth under the lid on all even-numbered trials from 1-20, but not on any other trial.

Extinction. An S who gave 10 consecutive responses of "out" following the last appearance of the cloth in the well was considered extinguished. Work with S was discontinued after that criterion was met.

Results. The mean number of responses of "in" during extinction-training and the mean number of trials to extinction are given in Table I. A response of "in" means that S guessed the red velvet cloth to be in the well of the apparatus.

As degree of training prior to extinction increased, both trial-scores and responses of "in" during extinction decreased. Bartlett's test for homogeneity permitted an analysis of variance for both measures. The variance due to amount of training was significant for trials ($F = 3.48$ for 2 and 42 *df.*, $P = 0.05$), but somewhat short

TABLE I
MEAN TRIALS TO EXTINCTION AND MEAN NUMBER OF RESPONSES
OF 'IN' DURING EXTINCTION

	Group 1		Group 2		Group 3	
	M	SD	M	SD	M	SD
Trials	12.86	6.07	12.20	6.56	7.73	3.90
Responses of 'in'	7.13	3.87	6.20	2.13	4.53	2.14

of significance at the 5% level for the number of responses of "in" ($F = 3.01$ for 2 and 42 *df.*, required $F = 3.22$). The subsequent *t*-tests for trials to extinction showed Group 3 significantly different from Group 1 at the 1% level and from Group 2 at the 2% level. Group 1 did not differ significantly from Group 2. The *t*-test was not applied to the data for number of responses of "in." The results taken together clearly indicate least resistance to extinction in Group 3, the most highly trained group.

Group 3 learned the pattern of reinforcement better than the other groups and could therefore better discriminate the deviation from the pattern represented in the extinction-series. The better learning of Group 3 may be illustrated by comparing the last acquisition-trial of this group with the last acquisition-trial of Group 2 and of Group 1. Since 15 Ss were used in each group, the maximal number of correct responses possible for each group on this trial was 15, a score of 7.5 correct indicating chance-accuracy. Group 1 had a score of only 6, which is below chance, while Group 2 made 8 correct responses, or just about chance. The members of Group 3 made 14 correct responses, a figure considerably above chance.

Summary. Three groups of college students were given different amounts of alternating partial reinforcement on a modified Wisconsin apparatus. The most highly trained group showed the least resistance to extinction. The results are in accord with those obtained in a previous experiment on rats in a runway. They are interpreted as evidence that S learns about the pattern of events in training. The greater the experience with the training pattern, the more readily does S discriminate the transition from training to extinction, and the greater the rate of extinction.

APPARATUS

CONSTANT-CURRENT SINE-WAVE CIRCUITS FOR STIMULATION

By HENRY E. GUTTMANN, University of Washington

A common feature of shocking circuits is a high external resistance in series with S , designed to minimize fluctuations in contact and tissue-impedance. It is not generally appreciated, however, that tissue-impedance varies inversely with the impressed voltage, even when an alternating sine-wave potential is employed to reduce the effects of polarization.¹ Duyff, who has made extensive measurements of non-nervous tissue-impedance *in vitro*, concludes that, in addition to the effect of voltage, tissue-impedance also is influenced by the frequency of the stimulating current, the temperature of the tissue, and the duration and pattern of successive stimulations.² Where the brain is stimulated directly by means of implanted electrodes, it is helpful, in evaluating changes in behavior that might be attributed to the formation of scar-tissue at the electrodes, to measure the impedance at the electrodes.³ In view of the large number of variables that may affect this impedance, it appears desirable to use an A.C. stimulating system that provides a 'constant' current, *i.e.* one minimally affected by the 'load' of the tissue being stimulated. Such a circuit has been described by Montgomery and Ward.⁴ Their circuit has the disadvantage, however, of being limited to line-frequency operation, and it requires a well-shielded power-transformer of a type not immediately available. The apparatus described here is based on the circuit of Montgomery and Ward, but it has been modified to permit full current output from 20-5000~. The shielded transformer has been eliminated; all components are commonly available radio replacement parts.

The circuit utilizes the high plate-resistance of pentode tubes to maintain

* The author is indebted to Dr. Roger B. Loucks for advice in the construction of this apparatus. He also thanks Mr. Bert Browne of the Speech Department and Professor W. Ryland Hill of the Electrical Engineering Department for many helpful suggestions.

¹ The use of direct current in the stimulation of the nervous system tends to cause the formation of gas bubbles when an electrode is enclosed beneath the skin, and this may cause injury to the tissue being stimulated.

² J. W. Duyff, *Low Frequency Impedance of Animal Tissue*, 1942, 1-108.

³ R. B. Loucks, Methods of isolating stimulation effects with implanted barriers, Chapter 12 in D. E. Sheer (ed.), *Electrical Stimulation of the Brain*, in press.

⁴ L. H. Montgomery and J. W. Ward, Constant current stimulator, *Electronics*, 24, 1951, 230-242.

a relatively constant plate-current flow independently of variations in the impedance of the tissue being stimulated. Two pentodes, with individual power-supplies, are used in a push-pull arrangement, each delivering one half of a sine-wave to the load. The circuit diagram (Fig. 1) shows that

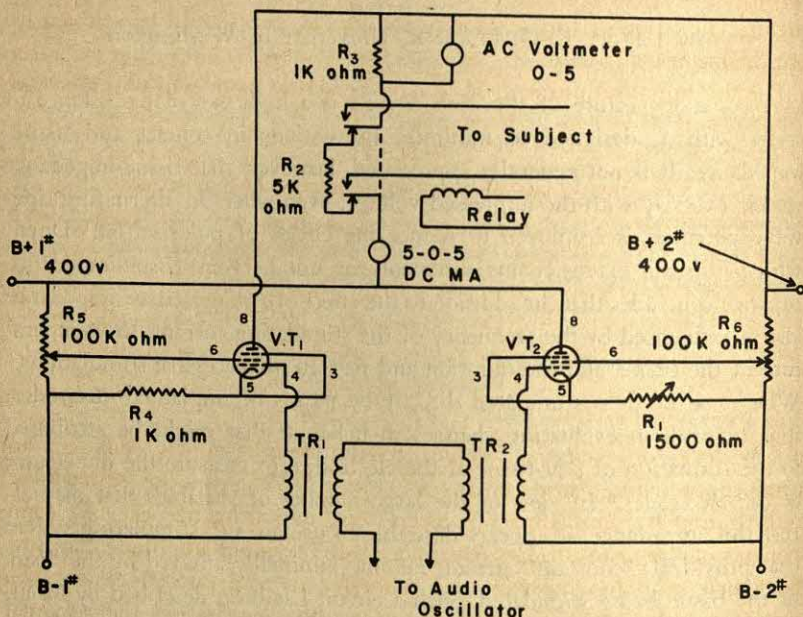


FIG. 1. CONSTANT-CURRENT SINE-WAVE STIMULATING CIRCUIT
 Relay, DPDT, make-before-break type; TR_1 , TR_2 , line-to-grid input-transformer;
 VT_1 and VT_2 are 6SJ7.

when VT_1 conducts, current flows through the load from power-supply 1, and, when VT_2 conducts, current flows through the load from power-supply 2 in the opposite direction. The negative terminals of the power-supplies are not brought to a common ground, because that would prevent proper functioning of the circuit. Two separate input-transformers are required for the same reason. Standard line-to-grid transformers are not mandatory; a large variety of small audio-transformers which were tested functioned satisfactorily provided that they were electrically similar. The output-voltages of the two power-supplies do not have to be identical, but the screen-voltages should be adjusted for equal voltages of approximately 125 v.

Adjustment and calibration. The 5000 Ω resistor (R_2) is used as a dummy load for adjustment purposes, and it is always in the circuit except when

stimulation is being applied to S . Since Relay 1 is of the make-before-break type, the dummy load is not disengaged from the circuit until S is fully contacted, in order to prevent a surge of current at the instant of contact. The 1000Ω resistor (R_3) provides a voltage-drop for the operation of an A.C. voltmeter calibrated to register the amount of current flowing in the circuit. The value of this resistor is not critical, and should be selected to suit the meter being used.

For the initial adjustment, R_1 should be set at about mid-position, and

TABLE I
VARIATIONS IN LOAD-CURRENT WITH DIFFERENT LOAD-RESISTORS

Substitute-load (K Ω)	Current (ma.)	Difference (ma.)	%
1	2.25	0.01	0.44
10	2.23	0.01	0.44
20	2.22	0.02	0.88
40	2.16	0.08	3.57
60	2.11	0.13	5.80
80	2.08	0.16	7.14
100	2.00	0.24	10.71

a low level of excitation applied to the input-terminals, sufficient to produce a reliable reading on the A.C. meter. R_1 is now adjusted until the null-reading D.C. milliammeter registers zero, indicating that the D.C. components of the plate-currents have cancelled. The input-amplitude may now be varied, and the A.C. voltmeter calibrated to read directly in ma. The wave-form should be observed on an oscilloscope to find the maximal current that can be drawn without distortion of the sine-wave. With the tubes and circuit-values specified, the input-voltage could be advanced to the level where 10 ma. were passed through the load before distortion of the wave-form occurred. The output and regulation of this stimulator are quite uniform from the lowest frequency tested (20~) up to 5000~. Above this frequency, the amplitude and regulation both decline appreciably.

Table I indicates the constancy of the output-current with differing load-values. All measurements were made with a 200~ input supplied by a Hewlett-Packard audio-oscillator (model 200-AB). Voltage-readings were taken across resistor R_3 , using a vacuum-tube voltmeter, and converted to milliamperes. Current-level was adjusted to an initial value of 2.24 ma. through a 5000Ω resistor at R_2 . The value of 2.24 ma. used for this series of measurements was arbitrary, and the data are submitted merely as representative.

As may be seen from Table I, regulation is at the 1% level for variations

(R) is adjusted to the point where breakdown occurs, *i.e.* the neon lamp (L) lights up. The potentiometer setting is then reduced slightly, and the reset switch (S_2) is opened. Upon the release of this switch (S_2), the voltage appearing across L is just below the breakdown voltage. At this time, any voltage induced in the secondary of the microphone transformer (T) will be in series with the 90-v. supply; thus any sound of sufficient intensity to induce the required differential voltage will trigger the lamp. The resulting flow of current energizes the relay. Lamp and relay both remain energized until the circuit is reset. Operation of this circuit is 'all or none,' and is completely free of chatter in the relay.

Since there is no amplification in this circuit other than that afforded by the input transformer, it is relatively insensitive, but it is adequate when the microphone is placed within a foot or so of a subject speaking in the microphone (M) in an ordinary tone of voice. If greater sensitivity is desired, two input transformers may be used with their secondaries connected in series-aiding. In our particular applications, however, the lack of sensitivity was not a disadvantage because it was found, when using high gain amplifiers, that the sensitivity had to be reduced, to a corresponding level, to eliminate responses to extraneous sounds.

AN INEXPENSIVE RESPIRATOR FOR USE WITH SMALL ANIMALS

By MATTHEW WAYNER, Syracuse University

In various experiments in the physiological psychology laboratory, *e.g.* those demonstrating the effects of curare and of certain spinal reflexes, artificial respiration is required. Because of their high cost, commercial respirators cannot be had, except in the wealthiest laboratories, in sufficient numbers to supply all the students. In addition to expense, the available respirators have two disadvantages: (1) the pump is close to the animal in order to make the dead-air space between the trachea and source of fresh air as small as possible; and (2) they apply pressure mechanically to the thorax for expiration and thus isolate too much of the animal from the experimenter—an important consideration when small animals are used. Fortunately, for students' use and for many research purposes, precise control of the amount of air or gas delivered to the animal's lungs is not important, only an adequate amount is necessary.

The respirator described here, which was used with the rat, is inexpensive—it can be made in a small laboratory—and it avoids both disadvantages mentioned above. It consists of two parts: a tracheal cannula, and an air pump. The cannula, pictured diagrammatically in Fig. 1, is made of two short pieces of 3/16-in. lucite

rods which are bored, milled and bonded together, as shown in the illustration, by lucite dissolved in chloroform. The tapered end of the cannula is so grooved that it will hold fast when tied after it has been inserted in the trachea. The opposite end of this tube is capped with a lucite disk through the center of which there is a small hole. The amount of air delivered to the animal's lungs is regulated by the size of this hole; the larger the hole the more air escapes through it and the less enters the lungs. A hole 0.02 in. in diameter is adequate for a pulse of air of approximately 5 cc. A series of interchangeable caps, each containing a hole of different size, permits the experimenter to vary rapidly the volume of the pulse and thus to obtain the amount desired. The animal exhales between successive strokes of the pump by means of the elasticity of its respiratory system, the expired air passing out of the cannula through the hole in the cap.

The angled arm of the cannula is attached by small rubber tubing to the air pump, a 20-cc. glass syringe. Between it and the tubing two one-way valves must be so inserted so that the successive pulses of air forced through the system will be

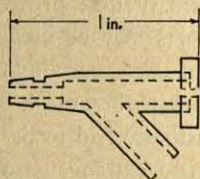


FIG. 1. TRACHEAL CANNULA

drawn from outside the system.¹ The pump may be operated manually, or mechanically by a simple device connected to a synchronous motor.

The apparatus is small; it does not isolate much of the animal from the experimenter, and it eliminates the problem of the dead-air space which requires that the pump be close to the animal. Occasionally mucous will block the tracheal end of the cannula. When this occurs it can easily be removed by pushing a fine wire through the hole in the cannula cap.

¹ Manometer bulb valves will work reliably if they are mounted in a vertical position that the operation of the rubber stops is facilitated by gravity.

NOTES AND DISCUSSIONS

AN HYPOTHESIS OF PSYCHOLOGICAL AGING

The hypothesis presented here is offered not as one derived from objective research but rather as one formulated from an oft-recurring awareness that human beings are nearly unanimous in their experience that chronological time passes ever more rapidly with advancing age. The experience of accelerating time is often explained either as the concomitant to a like increase of mental activity, or as a tendency on the part of aging persons to assign a greater specific value to life as an ever increasing fraction of it expires. Though such factors may play a part in the vague feeling held by different individuals of the same chronological age that their lives—as integrated wholes—have been 'long' or 'short,' they are not, as the present author believes, the *basic causes* of the awareness of *accelerating* time. In opposition to the supposed effect of increasing mental activity stands the observation that awareness of accelerating time is shared, for example, by laborers whose pattern of life has remained almost invariant since childhood. In opposition to the supposed effect of the increasing value of life stands the observation that awareness of accelerating time is shared by individuals who admittedly are disinterested in life and whose patterns of living fail in other respects to reveal any such metamorphosis of values.

It is the principal contention of the present hypothesis that psychological time bears a non-linear relationship to chronological time; that the duration of an interval in psychological time is subjectively measured by comparison with the total apparent duration of life up to a specific reference-date, which itself must be taken as an interval of psychological time. This hypothesis may be phrased more rigorously as follows: The apparent psychological duration of a chronological interval, compared with that of a second chronologically identical interval, is inversely proportional to the total apparent psychological duration of life up to and including the first interval compared with the apparent duration up to and including the second interval.

The following mathematical statement of this hypothesis allows one more easily to evaluate its significance:

Let the chronological age of a person in years be designated as n . The real duration of any year, i , of this person's life may be designated as $t_i =$

$\bar{t}$ = a constant number. The apparent (psychological) duration of the same year, which in general is not a constant, may be designated as τ_i . Now the real (chronological) duration of life of the person in n years is:

$$A_n = \sum_{i=1}^n t_i = n\bar{t}, \quad [1]$$

while the apparent (psychological) duration of life in the same period of n years is:

$$a_n = \sum_{i=1}^n \tau_i. \quad [2]$$

According to the foregoing hypothesis, we may state that

$$\tau_n = \frac{\sum_{i=1}^{n-1} \tau_i}{\sum_{i=1}^n \tau_i} \tau_{n-1} = \left(\frac{a_{n-1}}{a_{n-1} + \tau_n} \right) \tau_{n-1}, \quad [3]$$

which is equivalent to the standard equation

$$\tau_n^2 + a_{n-1}\tau_n - a_{n-1}\tau_{n-1} = 0. \quad [4]$$

Of the two possible solutions to Equation [4], only one is physically pertinent to the problem at hand. This solution is:

$$\tau_n = \frac{1}{2}(\sqrt{a_{n-1}^2 + 4a_{n-1}\tau_{n-1}} - a_{n-1}). \quad [5]$$

The simplest approach in evaluating Equation [5] is to set $\tau_1 = 1$ and solve for τ_n (n successively equals 2, 3, 4, ... , N) by the aid of Equation [2] in which $n - 1$ is substituted for n . Cumulative sums of τ_n computed from Equation [5], plotted as a function of n itself, yield the curve of Fig. 1.

The same data are presented in another manner in Table I. In this table may be found the chronological age at which an individual reaches the psychological mid-point of his life for the various decades to 100 years.

In accordance with the same hypothesis, Table II alternatively gives a comparison of the total apparent duration of each decade of human life as a decimal fraction of the apparent duration of various reference decades.

At least on the basis of the author's reflections on his own life, the relationship between chronological and psychological age shown in Fig. 1 and in Tables I and II appears to be reasonable. It is likely that moderate

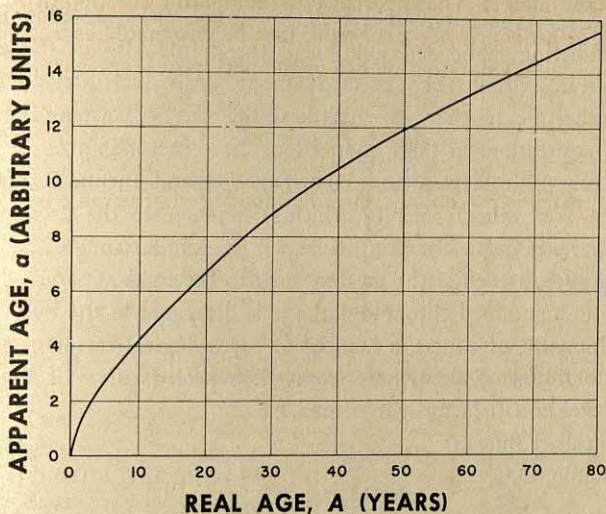


FIG. 1. RELATIONSHIP BETWEEN CHRONOLOGICAL AND PSYCHOLOGICAL AGE

TABLE I

AGE AT WHICH PSYCHOLOGICAL DURATION OF LIFE REACHES ITS MID-POINT

Total age (yr.)	Mid-point		Total age (yr.)	Mid-point	
	yr.	% total		yr.	% total
10	3.5	35	60	19.7	33
20	7.0	35	70	22.5	32
30	10.4	35	80	24.9	31
40	13.6	34	90	27.2	30
50	16.7	33	100	29.8	30

TABLE II

RELATIVE APPARENT PSYCHOLOGICAL DURATION OF EACH
CHRONOLOGICAL DECADE OF LIFE

Decade (yr.)	Decade of reference (in yr.)							
	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80
1-10	1.00	1.62	2.11	2.50	2.85	3.18	3.45	3.78
11-20	.62	1.00	1.30	1.54	1.76	1.96	2.13	2.33
21-30	.47	.77	1.00	1.18	1.35	1.51	1.64	1.79
31-40	.40	.65	.85	1.00	1.14	1.27	1.38	1.51
41-50	.35	.57	.74	.88	1.00	1.11	1.21	1.32
51-60	.32	.51	.67	.79	.90	1.00	1.09	1.19
61-70	.29	.47	.61	.73	.83	.92	1.00	1.10
71-80	.26	.43	.56	.66	.76	.84	.92	1.00

departures from the curve of Fig. 1 are to be expected among older persons following senile deterioration of the mind. Also it is likely that such departures are to be expected among individuals developing insanity or other mental changes which have the effect of varying the degree of mental occupation. In these connections, however, it is probably necessary to distinguish between mental *occupation* and mental *productivity or efficiency*, only the former of which is germane to the present issue; *i.e.* the awareness of accelerating time will be similar in the genius and in the moron even if the awareness as to the total apparent duration of life up to a specified chronological age is divergent in the two mentalities. Additionally, of course, a number of psychological factors can probably lead to *transitory* departures from the expected curve of Fig. 1—if these can ever be satisfactorily measured.

U. S. Weather Bureau
Washington, D.C.

J. MURRAY MITCHELL, JR.

DARK-ADAPTATION AND CONDITIONING: SOME OBSERVED CORRELATIONS

In view of current attempts to integrate the fields of personality, learning, and perception,¹ a note on some correlations observed by the authors during recent experiments on dark-adaptation and conditioning may be of some interest. The Ss were students (men) in fields other than psychology. The measures of dark-adaptation represent the times taken to detect a parafoveal test-field, presented at six different luminance-levels, following a period of light-adaptation. The luminance-levels, in log units (μ L.), are shown in the first column of Table I. All times were measured from the beginning of dark-adaptation. Full details of apparatus and procedure are reported elsewhere.² The conditioned reflexes studied were the eyeblink and the galvanic skin response (GSR), the conditioning laboratory and pro-

* From the Institute of Psychiatry.

¹ H. J. Eysenck, Cortical inhibition, figural after-effect and theory of personality, *J. abnorm. soc. Psychol.*, 51, 1955, 95-105; C. M. Franks, Personality differences in the visual perception of verticality, *Rev. Psychol. Appl.*, in press; Conditioning and personality; a study of normal and neurotic subjects, *J. abnorm. soc. Psychol.*, 52, 1956, 143-150; Franks, Recidivism, psychopathy and personality, *Brit. J. Delinq.*, 6, 1956, 192-201.

² G. W. Granger, Dark adaptation in anxiety states and hysterics, *Brit. J. physiol. Opt.*, 13, 1956, 234-241.

cedure are described in detail elsewhere.³ Partial reinforcement was used, followed by a series of successive nonreinforced trials. This gave two convenient measures of the rate of conditioning, the number of CRs to the test-trials during acquisition interspersed among the reinforced trials and the number of CRs to the successive test-trials during extinction.

Table I shows the product-moment correlations between the six measures of adaptation and the scores of acquisition and extinction for both eyeblink and GSR. Correlations with conditioning of the eyeblink are based on the results from 46 Ss, those with the GSR from 36 Ss. These correlations, all of

TABLE I
CORRELATIONS BETWEEN DARK-ADAPTATION AND CONDITIONING

Luminance (log μ L.)	Eyelid		GSR	
	acquisition	extinction	acquisition	extinction
(1) 6.33	.278	.405†	.364*	.337*
(2) 5.94	.323*	.481†	.395*	.466†
(3) 5.53	.135	.257	.059	.130
(4) 4.96	.087	.236	.111	.050
(5) 4.57	.151	.343*	.160	.081
(6) 4.16	.232	.323*	.322	.152

* Significant at 5% level. † Significant at 1% level.

which are positive, indicate a tendency for long perceptual times to be associated with ease of conditioning and resistance to extinction. Although none of the correlations is high, several are statistically significant, particularly for perceptual times at the two highest levels of luminance. These perceptual times represent the early stages of dark-adaptation in the first few minutes following light-adaptation, at what are conventionally regarded as 'photopic' (*i.e.* above about 6 log unit μ L.) rather than 'scotopic' levels of illumination.

Interpretation of the coefficients of correlation is often difficult but it is particularly so in the present case where there is hardly any previous research to guide us. Certainly there are no grounds for supposing that conditioning and dark-adaptation depend upon the same underlying mechanisms. Whereas dark-adaptation seems to depend primarily upon peripheral mechanisms in the retina, conditioning occurs within the

³ Franks, The establishment of a conditioning laboratory for the investigation of personality and cortical functioning, *Nature*, 175, 1955, 984-985; Franks, *op. cit.*, *J. abnorm. soc. Psychol.*, 52, 1956, 143-150; Effect of food, drink, and tobacco deprivation on the conditioning of the eyeblink responses, *J. exp. Psychol.*, 1957, 117-120.

central nervous system (CNS). Even if, as some authors claim,⁴ there is a neural as well as a photochemical component in visual adaptation, the neural mechanisms involved still appear to be retinal in origin.

Even though dark-adaptation under normal conditions appears to be largely determined in the retina, the actual value obtained for the threshold-energy in a dark-adaptation experiment is of course dependent upon the state of the conducting pathways between the retina and the brain, and the receiving areas in the cortex. For instance, if *S* is suffering from anoxia, the visual threshold may be raised even though anoxia does not apparently interfere with the mechanism of dark-adaptation *per se*.⁵ Now if conditioning is also dependent upon the excitability of the CNS, it is quite conceivable that insofar as both visual thresholds and CRs reflect individual differences in central excitability then a basis for correlation between them may be found.

A possible lead for further research is provided by the results of studies with drugs. There is some evidence that certain depressants such as alcohol impair conditioning⁶ and also raise light-thresholds.⁷ Owing to marked differences in experimental conditions, it is difficult to determine the implications of this research for the present study. It seems possible, however, that certain depressants of the CNS would tend to reduce the effective intensity of previous light-adaptation and thereby facilitate the recovery of sensitivity during subsequent dark-adaptation. This assumes that the drug exerts its effect only during the process of light-adaptation. If it continued to have an effect during dark-adaptation, the outcome would be very difficult to predict. A further difficulty in interpretation stems from the lack of evidence regarding the existence of conditionability as a general property of the CNS. This makes it premature to generalize about possible correlations between measures of dark-adaptation and measures of conditioning to reflexes other than the eyeblink and GSR. It is obvious that more data are needed concerning both the generality of conditioning and the actions of drugs before useful hypotheses can be developed.

University of London

G. W. GRANGER

C. M. FRANKS

⁴R. J. Lythgoe, The mechanism of dark adaptation, *Brit. J. Ophthalm.*, 24, 1940, 21-43.

⁵R. A. McFarland and J. N. Evans, Alterations in dark adaptation under reduced oxygen tensions, *Amer. J. Physiol.*, 127, 1939, 37-50.

⁶N. A. Popov and C. Popov, Contribution à l'étude des fonctions corticales chez l'homme, par la méthode des réflexes conditionnés électrocorticaux: Action de l'alcool sur les images consécutives et leur conditionnement, 1^o note, *C. R. Acad. Sci. Paris*, 237, 1953, 930-932.

⁷J. Mandelbaum, Dark adaptation: some physiologic and clinical considerations, *Arch. Ophthalm.*, 26, 1941, 203-239.

KAREN'S STUDY OF INCIDENTAL LEARNING AS A
FUNCTION OF MEANINGFULNESS

In a recent issue of this JOURNAL, Karen reported an experiment concerned with the difference between intentional and incidental learners as a function of the meaningfulness of the learning materials.¹ He concluded that his results were at variance with earlier findings by Postman, Adams, and Phillips.² The purpose of this note is to comment on Karen's procedure and to examine the validity of his conclusion.

In the experiment of Postman, Adams, and Phillips, the learning materials consisted of a series of 20 nonsense-syllables varying widely in association-value. The items were presented on a memory drum at a 4-sec. rate. The Ss were run in pairs. The incidental learners served as 'experimenters' and intentional learners served as 'subjects.' After one presentation of the list, retention was tested either by the method of free recall or by the method of recognition. Intentional Ss surpassed incidental Ss in the free recall of items of low association-value but not of high association-value. There were no differences between the two kinds of learners on the test of recognition.

Karen's materials consisted of 30 items, half of which were nonsense-syllables and half meaningful words. The items were presented on cards at a 15-sec. rate. As a given item was exposed, the incidental Ss had the task of finding the position of that item in a matrix which contained *all the items in the learning list*. Intentional Ss performed the same task and were, in addition, instructed to memorize the materials. After one presentation of the series, a test of recognition was given. The recognition-scores of the intentional learners were higher than those of the incidental learners, and the difference between the two kinds of learners was significantly greater for meaningful words than for nonsense-syllables. These results were presented as an apparent contradiction of the findings of Postman, Adams, and Phillips.

The results of the two experiments are not comparable for two major reasons: (a) Karen's failure to exercise reasonable control over the exposure of the learning materials, and (b) a critical difference between the orienting tasks used in the two experiments.

As was indicated above, the total series of learning items was available

¹ R. L. Karen, Recognition as a function of meaningfulness and intention to learn, this JOURNAL, 69, 1956, 650-652.

² Leo Postman, P. A. Adams, and L. W. Phillips, Studies in incidental learning: II. The effects of association value and of the method of testing, *J. exp. Psychol.*, 49, 1955, 1-10.

to Karen's Ss throughout the training period. As a result, there was no effective control over the amount of time spent on individual items or classes of items. This fact in and of itself vitiates any conclusions drawn from Karen's data. In addition, the simultaneous presence of the total list provided Ss with considerable opportunity for rehearsal. It is reasonable to assume that intentional Ss made more extensive use of this opportunity than incidental Ss.³ Meaningful words are, of course, acquired at a faster rate than are nonsense-syllables. Hence, meaningful words would derive a differential advantage from uncontrolled rehearsal. Thus, the failure to control opportunities for rehearsal biases the results in favor of superior retention of the meaningful words by the intentional Ss.

In asking his Ss to locate the position of each item in a matrix comprising the total learning series, Karen used an orienting task highly conducive to intraserial interference. Repeated hurried perusals of the list clearly favor generalization among the items. Intraserial interference would develop at a faster rate in intentional than in incidental learners. In attempting to memorize the entire series, intentional Ss are very likely to form more incorrect associations than incidental Ss. Such interference effects would, of course, be more serious for nonsense-syllables than for words. Thus, the conditions of interference again favor superior retention of the words by intentional learners.

In an analysis of the effects of different orienting tasks, Postman and Adams have shown that procedures conducive to intraserial interference may, indeed, significantly change the effect of meaningfulness on the difference between intentional and incidental learners.⁴ With an orienting task specifically designed to produce intraserial interference, they found a difference in favor of intentional learners with meaningful words but not with nonsense-syllables. The interference from the orienting task had, as it were, shifted the scale of difficulty of the learning items. Recall of the nonsense syllables was uniformly low, and meaningful words were of a difficulty adequate to discriminate between intentional and incidental learners. Thus, Karen's results could be accounted for in terms of the conditions of interference inherent in his experimental procedure. His failure to control the conditions of exposure and rehearsal, however, makes any interpretation of the results highly tenuous.

University of California

LEO POSTMAN

³ Karen does not report a post-experimental inquiry to determine whether any of the incidental Ss had adopted a set to learn.

⁴ Postman and Adams, *Studies in incidental learning: IV. The interaction of orienting tasks and stimulus materials. J. exp. Psychol.*, 51, 1946, 329-333.

INDEX OF MEANS VS. MEAN OF INDICES

Recently the writer read a study in which the author had estimated the mean IQ of students in each of several grades in an elementary school by dividing the mean mental age ($\overline{MA}$) of each group by its mean chronological age ($\overline{CA}$) and multiplying this quotient by 100.

The above procedure may seem to be an ingenious short-cut when $\overline{MA}$ and $\overline{CA}$ have already been computed. For many years, however, statisticians have known that, in general, the index of means, $\overline{X/Y}$, is not equal to the mean of indices,

$$\sum_{i=1}^N (X_i/Y_i)/N.$$

They *may* be equal but are not necessarily so.

Rao gives a formula from which we can infer that $\overline{X/Y} = \overline{\overline{X}/\overline{Y}}$ when $\overline{X/Y}$ equals the regression coefficient $r_{xy} s_x/s_y$.¹

A different approach, devised by the present writer and outlined below, shows rather simply that only if the indices X_i/Y_i do not covary with their denominator, Y_i , will $\overline{X/Y} = \overline{\overline{X}/\overline{Y}}$, and that under typical school conditions $100(\overline{MA}/\overline{CA})$ underestimates the mean IQ .

Proof. We shall consider first quotients more general than IQ s, utilizing expected values. Let X and Y denote any two real-valued variates. We want to determine under what conditions the following equality holds:

$$E(X)/E(Y) = E(X/Y). \quad [1]$$

Multiplying both sides of Equation [1] by $E(Y)$ yields

$$E(X) = E(X/Y)E(Y). \quad [2]$$

Now, the formula for the product-moment coefficient of correlation between X/Y and Y may be written as follows:

$$\rho_{(X/Y),Y} = \{E[(X/Y)Y] - E[(X/Y)E(Y)]\}/\sigma_{(X/Y)}\sigma_Y \quad [3]$$

Therefore,

$$-\rho_{(X/Y),Y}\sigma_{(X/Y)}\sigma_Y + E(X) = E(X/Y)E(Y). \quad [4]$$

Substituting the left-hand side of Equation [4] for the right-hand side of Equation [2] and simplifying, we obtain the covariance of

¹ C. R. Rao, *Advanced Statistical Methods in Biometric Research*, 1952, 154.

X/Y with Y equals zero,

$$\rho_{(X/Y), Y\sigma_{X/Y}\sigma_Y} = 0, \quad [5]$$

as the condition necessary for Equation [1] to hold.

Application to IQ. There are four possible situations, or combinations, in which $100(\overline{MA}/\overline{CA}) = \overline{IQ}$: $N=1$, $s_{IQ}=0$, $s_{CA}=0$, or $r_{IQ,CA}=0$.

Only the last is of practical importance, since almost never will the standard deviation of IQs or CAs be zero. If $r < 0$, then $100(\overline{MA}/\overline{CA}) < \overline{IQ}$. If $r > 0$, then $100(\overline{MA}/\overline{CA}) > \overline{IQ}$.

For an unselected population and a properly standardized test, $E(MA) = E(CA)$ and $E(IQ) = 100$, so $\rho_{IQ,CA} = 0$. Because of late entrance to school, non-promotion, and acceleration, $r_{IQ,CA}$ within a grade is, however, usually negative.² In other words, the brighter youngsters tend to be younger, the duller ones older.

Therefore, $100(\overline{MA}/\overline{CA})$ tends to underestimate $\overline{IQ}$. The extent of this underestimation varies from school to school, grade to grade, and test to test.

University of Wisconsin

JULIAN C. STANLEY

TWENTY-EIGHTH ANNUAL MEETING OF THE EASTERN PSYCHOLOGICAL ASSOCIATION

The Eastern Psychological Association met April 12 and 13, 1957 at the Hotel Statler in New York City. Approximately 2300 persons registered at the meetings. Of these, 1,458 were members of the Association, 522 were guests, and 320 were new members who joined the Association at the meetings. The present active membership of the Association totals 3,084.

A committee headed by Joseph M. Notterman was responsible for local arrangements and the program was planned by a committee under the chairmanship of James Deese. The program consisted of 195 papers (presented in 28 sessions), 1 symposium, 7 special meetings, 3 invited addresses, and special films by McGraw-Hill. The 28 sessions were concerned with the following topics: vision (2 sessions), clinical processes, human learning (2 sessions), social psychology (2 sessions), brain function (2 sessions), animal behavior (3 sessions), physiological (2 sessions), group processes, applied psychology, sensory (nonvisual), clinical testing,

² Quinn McNemar, *Psychological Statistics*, 2nd ed., 1955, 163.

verbal behavior, aversive conditioning, visual perception, industrial psychology, clinical psychology, stress, operant behavior, probability learning, personality, and developmental. The symposium was titled: "Problems, experiments and issues in the transfer effect of early exposure on perceptual discrimination." In addition to special meetings entitled "Group dynamics and group psychotherapy" and "Autistic and schizophrenic children," the program included a meeting of Psi Chi and meetings concerned with the cooperative psychological research project in tuberculosis, American Association of University practicum training centers, conference of state psychological associations, and the licensing procedures in New York. The following invited addresses were presented: "Psychology and the design of machines," by Franklin V. Taylor; "Perception and perceptual learning," by Donald O. Hebb; and "Soviet psychology since 1950," by Gregory Razran. Special films were shown by McGraw-Hill.

Fred S. Keller presented the annual presidential address, "The phantom plateau." During the business meeting it was announced that the following new officers had been elected: President, Stuart W. Cook, and Directors, Lorrin Riggs and E. Joseph Shoben, Jr.

The 1958 meetings of the Association will be held at the Bellevue-Stratford in Philadelphia, April 11 and 12. The 1959 meetings will be held in Atlantic City.

University of Delaware

GORHAM LANE

TWENTY-NINTH ANNUAL MEETING OF THE MIDWESTERN PSYCHOLOGICAL ASSOCIATION

The Midwestern Psychological Association met at the Hotel Sherman, Chicago, May 2, 3, and 4, 1957. Northwestern University and the University of Chicago were the host institutions. George S. Speer, Illinois Institute of Technology, served as Chairman of the Local Arrangements Committee. Registration for the meeting was 1,220. The program consisted of 117 papers, 10 symposia, and 3 seminars. The symposia were concerned with theoretical interpretations of runway studies of partial reinforcement, psychological effects of cerebral lesions in human beings, a critical evaluation of the effectiveness of the team approach in psychiatric outpatient clinics, the development of processes for indirect or synthetic validity, psychology as a basic science in medical education, challenges and accomplishments of management development in business and in-

dustry, stimulus satiation, evaluating psychotherapy, the acquisition and transfer of skills, and industrial morale. The presidential address, "Of models and men," was delivered by William K. Estes of Indiana University.

Delos D. Wickens, Ohio State University, was elected president for 1957-58. Cletus J. Burke was elected to the Council for 1957-60. With the election of 268 new members, the current membership is 2082. The 1958 meeting will be held May 1, 2, and 3 at the Hotel Sheraton-Cadillac, Detroit.

University of Chicago

DONALD W. FISKE

FOURTH INTERAMERICAN CONGRESS OF PSYCHOLOGY

The Fourth Interamerican Congress of Psychology, sponsored by the Interamerican Society of Psychology, was held at the University of Puerto Rico, Rio Piedras, P.R., from December 26 to 30, 1956. The central theme of the Congress, "The psychology and psychopathology of group behavior," was discussed in a series of nine symposia (based on the presentation of 64 research reports and surveys), as follows: "National characteristics of behavior: general," "National characteristics of behavior: Latin American," "Group interaction," "The anti-social behavior of groups," "Interpersonal relations," "Childhood and adolescence," "Group psychotherapy," "Religion and group behavior," and "Experiments and tests."

The University of Puerto Rico and the Puerto Rican Psychological Association were co-sponsors of the Congress. A special evening reception, among other social events, was given to all delegates at La Fortaleza, San Juan, by Governor and Mrs. Muñoz Marín. Delegates to the Congress were from Brazil, Canada, Colombia, Mexico, Puerto Rico, the United States, and Venezuela.

The officers elected for 1957 at the business meeting of the Congress were: President, Dr. Otto Klineberg, Columbia University, U.S.A.; President-elect, Dr. Guillermo Davila, National University, Mexico; vice-President, Dr. Gustave M. Gilbert, Michigan State University, U.S.A.; Secretary-General, Dr. Werner Wolff, Bard College, U.S.A.; Executive Secretary, Dr. Samuel Pearlman, Brooklyn College, U.S.A.; Treasurer, Dr. Robert B. Malmø, McGill University, Canada.

An initial distribution of a published Symposium of the Third Interamerican Congress of Psychology (Texas, 1955), entitled "Psychological approaches to intergroup and international understanding," was made.

Efforts will be made in due time to publish the papers of the Fourth Congress.

The 1957 meeting of the Society will be held, according to present plans, at the National University of Mexico, Mexico City, in December.

Bard College

WERNER WOLFF

ERRATA

R. W. Moncrieff, the author of "Olfactory adaptation and odor intensity" in the March, 1957 number of this JOURNAL (Vol. 70, pp. 1-20), pointed out in a recent communication that two errors appear in this article. One occurs on Page 11, third line from the bottom, where "1.0" should read "10"; the other on Page 13, thirteenth line from the bottom, where "undiluted" should read "10%."

Thomas Mulholland, the author of "The frequency procedure for analyzing the data obtained by the method of limits" in the June, 1957 number (Vol. 70, pp. 280-282), pointed out in a recent communication that three errors appear in Table VI (Page 282). Under "SD" in the table, the figures, reading down, should be: "4.02," "5.20," and "4.42."

AN ACKNOWLEDGMENT

The JOURNAL is indebted to Professor Ernest R. Hilgard for supplying the photograph of Professor Lewis M. Terman, which appears in the frontispiece of this number. The photograph was taken in the spring of 1942, just before his retirement at the age of 65 years. The signature was taken from a letter received from him by the editor in April, 1955, when Professor Terman was 78 years old.

K. M. D.

Lewis Madison Terman: 1877-1956

Lewis Madison Terman, productive psychologist to the end, died on December 21, 1956, a few weeks before reaching the age of 80 years. He was born January 15, 1877, in Johnson County, Indiana, the twelfth of fourteen children in a farmer's family, the son of James William and Martha Cutsinger Terman. He died from a series of cerebral hemorrhages that had incapacitated him for a few weeks prior to his death, but even within a week of his death he was able to arouse himself enough to inquire about the topics discussed at a scientific meeting. He was last in his office on September 27, at work on a fifth volume in the series of studies on the intellectually gifted.

In his autobiography, prepared in 1931, when he was 54 years old,¹ he gave a very candid account of the formative period of his childhood and of the intellectual experiences of his youth. He reports Scotch-Irish, Welsh, German, and French descent, without any known ancestors who had achieved distinction. Previous to his own generation he knew of no one either on his father's or his mother's side who had belonged to any of the professions or had graduated from college.

Lewis entered a one-room school three months before his sixth birthday and at the end of the first term of six months was promoted to the third grade—an early evidence of unusual ability. Somewhat given to introspective reflection, he performed a memory experiment at the age of 11 years, with himself as subject. He remained in the country school for a year beyond graduation from the eighth grade at the age of 12 years. After work on the farm he entered the Central Normal College at Danville, Indiana, at 15 and remained for two years. With country-school teaching in between, he ended with both a B.S. and a B.Pd. (Bachelor of Pedagogy) in 1897, but took an extra year of training in 1897-1898. He now became principal of a township high school in his home county (Smiths Valley, Indiana), 1898-1901. With borrowed money he continued his education at Indiana University (A.B., 1902; A.M., 1903), where he came under the influence of Bryan, Lindley and Bergström. Then, with a fellowship and more borrowing, he went on to Clark University where he received the Ph.D. degree in 1905. Here he worked under Hall, Burnham and Sanford. Among his fellow-students were Book, Conradi, Gesell, Huey, Kuhlmann, Morse, and Porter.

Dr. Gesell writes of their common interest in the period of childhood,²

¹ L. M. Terman, *A History of Psychology in Autobiography*, Clark University Press, II, 1932, 297-331.

² Private communication from Arnold Gesell, April 18, 1957.

while at Clark University and how they both sensed the genius and personality of G. Stanley Hall. Speaking of Lewis Terman, Gesell tells a little anecdote which characterizes the freedom with which natural opportunities were capitalized upon for observation.

He was not perturbed by the almost forbidding immensity of the steep wooden staircase which gave monumental entrance to the main academic building. He was preoccupied with problems of stupidity and intelligence. So why not put this formidable incline to scientific use? The flight of steps could test the motor coordination of the bright and the backward boys who were under comparative study! Accordingly, Terman marshaled the boys and I served as time-keeper, recording the performance for speed of climbing and descent. Puzzle books were explored for the testing of more subtle intellectual processes—Binet would soon appear on the horizon.

The mere recital of training and degrees tells little of the determination that was required. Terman had married Anna Minton, a fellow student at Danville, in 1899. Already in 1899 and in 1900 there were suggestions of the undiagnosed tuberculosis that was to erupt into pulmonary hemorrhages in the summer of 1904, while between the two years at Clark. The patient care and understanding of his wife permitted him then to remain at his studies, despite her need to minister both to him and to the two young children now in the household.³

He prepared a manuscript describing his illness in 1953, at the age of 76 years, intending it for inclusion in a book of testimonials by ex-tubercular patients.⁴ In this he tells of the events and the recurrent anxieties. The tuberculosis determined the choice of a mild climate; a choice between Florida and California was decided in favor of the latter. But eight weeks after arriving for the new job of high school principal at San Bernadino in 1905 the old hemorrhage recurred. He had not admitted his tubercular history to his employers; instead of confessing it now, he rested and worked at home for a short while and returned to full duty eleven days after the hemorrhage. Fortunately there were no ill effects. In later years he suffered three attacks of pneumonia, and two flare-ups of tuberculosis (1928 and 1936), but none of these meant more than six weeks away from work. These recurrent illnesses gave him an interest in health, resulting in several books and papers on school hygiene, so that he was fond of quoting: "If you scratch a health reformer you will find an invalid."⁵

³ Mrs. Terman died a few months before her husband. The children, Frederick Emmons Terman, and Mrs. Helen Clare Mosher, both survive, and both gave him grandchildren.

⁴ *I Was Lucky*, 1953 (manuscript).

⁵ *Autobiography*, 324. It is of some interest that there was a history of tuberculosis in his family, for most of his father's nine siblings died of it before the age of 50 yr.,

After a year in San Bernadino Terman went as professor of psychology to the Los Angeles State Normal School to teach child study and pedagogy (1906-1910). There he was associated with Arnold Gesell, a friend from Clark days, and they lived as neighbors in San Fernando Valley (Valley View Road, Casa Verdugo), then a rural region, accessible to the college by trolley. Dr. Gesell writes that the Termans were good neighbors, that the boy Fred helped to construct the bungalow that Arnold and his brother, Robert, were building with their own hands, and that Lewis introduced Arnold to the fellow faculty member who later became Mrs. Gesell.

Terman left Los Angeles to accept a position as assistant professor of education at Stanford University (1910-1912), and there became successively associate professor (1912-1916) and professor (from 1916 on).

Stanford had had an adequate undergraduate department of psychology under Frank Angell, who had been brought there in 1892 from Cornell, where he was succeeded by Titchener. While the equipment was good and the instruction sound, graduate work had not developed—there had been only one Ph.D. granted in 30 years. Upon Angell's retirement President Wilbur did the unexpected thing of moving Lewis Terman from the School of Education to become head of the Department of Psychology. The 20 years (1922-1942) that Terman served as head marked the establishment of Stanford's department as a major graduate and research center. There were 51 Ph.D.s awarded in the years 1923-1942. Of the staff members he brought to Stanford, three (in addition to himself) became presidents of the American Psychological Association; of the students who were there during his tenure, four have become APA presidents. Thus eight presidents of the APA over the past 35 years were Stanford-connected during Terman's tenure as department head—no mean achievement in view of the many better-established departments at the time he was placed in charge. He was proud also of his "near successes" in staff recruitment, for he had made efforts to interest both Boring and Lashley in joining the staff when they were still in their early thirties.

Beyond his retirement he kept an office and staff at the University, and, with the help of foundation and government support, continued his research and writing. A number of health accidents occurred after his retirement—severe burns as a result of a fire in his residence, a broken hip,

his own eldest brother died of it in his childhood, and his own father had some hemorrhages while Lewis was a young man, although the father lived until 76. His son Frederick also suffered from tuberculosis in his early youth, recovering after two years of illness to continue in a distinguished and energetic career as an authority on electronics, Dean of the School of Engineering, and Provost of Stanford University.

operations for cataracts in both eyes—yet he continued to be cheerful, essentially youthful in outlook, and always forward-looking. He was as interested in the presidential elections of 1956 as he had been in earlier elections—and continued to write letters to radio commentators. His major concern continued to be with the gifted group—to whom he addressed general letters headed "To my 'gifted children.'" One such letter, mailed in January 1955, asked for additional information beyond a 1950-1952 follow-up, and expressed gratification at some of the current attention being given gifted children. "Our studies are bearing fruit—studies that would not have been possible without the extraordinary coöperation of my 'children' and their parents during the more than 30 years since our project was launched."

His scientific work brought him numerous honors, among them honorary doctoral degrees from the University of California, Indiana University, University of Pennsylvania, University of Southern California; an Honorary Fellowship in the British Psychological Society and an Honorary Fellowship of the Educational Institute of Scotland; presidency of the American Psychological Association, election to the National Academy of Sciences and the American Philosophical Society. He was invited to deliver the first Walter V. Bingham Memorial Lecture on March 25, 1954, when already 77 years old.⁶

CONTRIBUTIONS TO PSYCHOLOGY

There are several ways in which a man can contribute to his profession: as an investigator, by adding to its substance; as a teacher, by inspiring the young and enlisting able students into the ranks; as an administrator, by facilitating the institutional arrangements under which a science and profession can flourish. Professor Terman left an enviable record in all three of these.

Investigator. In the shadow of Stanley Hall it was no surprise to see Terman plan his researches in bold strokes and on a large scale. Chief among these are the revisions of the Binet-Simon scales that were known first as the Stanford-Binet (1916), and then as the Revised Stanford-Binet (with Maud A. Merrill) (1937), the Stanford Achievement Tests (with T. L. Kelley and G. M. Ruch) (1923), the series of studies of gifted children and their subsequent development, known as Genetic Studies of Genius (first volume 1925, fifth volume to appear posthumously under the joint authorship of Melita Oden), the development of a masculinity-femininity test and its use in characterizing sex and personality (with

⁶ Terman, The discovery and encouragement of exceptional talent, *Amer. Psychologist*, 9, 1954, 221-230.

Catherine C. Miles, 1936) and a major study of psychological factors in marital happiness (1938).

The large-scale methods he had used in test standardization prepared him well for his service with the Army in the First World War in the development of the Army Alpha and the Army Beta, where he served with the rank of major. There he gained the respect of his psychological colleagues, to many of whom intelligence-testing had previously been alien, associated with educational technology rather than with psychological science. So Terman became re-identified with psychology, served on the Council of the American Psychological Association from 1919 through 1921, and was elected its President in 1923.

His investigatory work was ingenious but always empirical. He noted in his autobiography that he had contributed little to theory. The tests he developed leaned heavily on Binet, for whom he had the deepest admiration. Although he made the *IQ* a term in the common vocabulary, the original idea of the *IQ* was not his but Stern's. His book on sex and personality did not cite Freud at all, and the one on marriage referred more frequently to Havelock Ellis than to Freud. To say that his work was empirical rather than theoretical does not mean that it was uncreative or uncritical. On the contrary, the good sense in the work set standards by which theoretical work could be tested.

The study of which he was most proud was that of the gifted group. His capacity to hold to a large-scale enterprise is shown by his persistence in following the 1500 bright children, of average age 11 years, first tested in 1922. He was continuously in touch with them, their spouses, children (and grandchildren!), so that, in 1957, 98 percent of those still living were in active participation in the study. The study will go on not only as a living monument to him, but as an increasingly valuable source of data on the life histories of able people. Today, when resources of talent are being treasured as never before, Terman's work stands as a landmark in the identification of talent and in proposing its more effective utilization.

Teacher. Terman's influence as a teacher is best evidenced by the many people of ability who worked with him, chose his type of problem, and identified themselves strongly with these interests throughout their later careers.⁷

An outstanding memory for them (and for other Stanford students of his day) is the Monday evening seminar in his home—reminiscent of

⁷ A list of his Ph.D. students up to 1931 can be found in his autobiography. Names of those strongly identified with psychological testing include, among others, A. S. Otis, S. C. Kohs, G. M. Ruch, M. A. Merrill, F. L. Goodenough, C. C. Miles, B. S. Burks, and R. G. Bernreuter.

Stanley Hall's—where, first of all, the invitation to attend was an honor, and where a high level of interchange took place. He sat toward the back of the room in an easy chair while a student presented his own experimental work, or reviewed a new book or a development within the field. He entered the discussion quietly, often warmly appreciative of an able job, but then raised a few trenchant questions that left no doubt of his skill, breadth of knowledge, and critical acumen. If the student was shoddy or in any sense bluffing, the ordeal could be a punishing one. Mrs. Terman usually invited a few of the students to dinner before the seminar, and both of the Termans soon were calling all of the graduate students by their first names, and taking a personal interest in their wives and children.

He valued originality, and was kinder to a student fumbling with a new idea than with one dealing meticulously but pedantically with a trivial one. His courses were not systematic—he disliked textbooks, and had a distaste for textbook writing—but he liked to work with ideas that kept psychology in a large setting and kept students asking questions and seeking their own answers to them. One of the courses he taught for a number of years, for example, was on the psychological analysis of biographies. He asked for no standard pattern: the student merely tried to make sense of the motives and achievements of a person significant enough to have a book written about him.

Even those who did not work in his own field felt strongly the influence of Terman's presence as an admirable identification figure. In considering teaching skill it is sometimes overlooked that a teacher is often greater for what he stands for as a person and as a scholar than for his techniques of instruction.

Administrator. Administration is sometimes dismissed as paper-pushing, whereas, at its best, it creates an atmosphere in which personal relationships are good and work of a high level gets accomplished. By these standards, Terman was a very successful administrator. In reviewing the history of Stanford's psychology department under his leadership, in a private statement to President Wilbur prepared at the time of his retirement, he wrote:

Trends and problems of any science change with the passing of time, but I think that certain basic traditions have been established here that will carry forward—recruiting of the ablest appointees available, the use of democratic administrative procedures, the encouragement of team work within the department, the fostering of close relationships with other departments, and the guaranty of academic freedom combined with personal responsibility.

Terman's administrative skill was exercised quietly and unobtrusively.

He came to his office at 10 A.M., spent a half-hour with his secretary, had a half-hour free for appointments with students or staff, and could not run over because he taught a class at 11:00. This was the only visible picture of his management of a growing and lively department. He did a certain amount of committee work in the afternoons, and was often called upon by President Wilbur for advice. He was largely responsible for Stanford's early leadership in using aptitude tests in selective admission, and he encouraged an independent study program that has continued to provide flexibility in undergraduate instruction.

SYSTEM OF VALUES

Because of Terman's championship of the intellectual elite, and his position on the side of heredity in the nature-nurture issue, it would be easy to jump to the mistaken conclusion that he was a social conservative. In fact, Pastore, testing the hypothesis that the broader social frame of reference was correlated with the position on the heredity-environment issue, found Terman an exception to his generalization.⁸

Pastore quotes a letter written to him by Terman in March, 1949, at the age of 71 years. In it, while disclaiming that he is a socialist, Terman states that he is not afraid of partial socialization as in education and medicine. "Most of all," he wrote, "I believe in civil liberties of the kind supposedly guaranteed by the Bill of Rights. Our failure to insure those rights to minority groups I consider a national disgrace. Nothing disturbs me more than our widespread racial and religious discrimination."

Apparently the hereditarian emphasis was itself weakening with the years. At the time of the autobiography (1932) he had written his belief: "That the major differences between children of high and low *IQ*, and the major differences in the intelligence test-scores of certain races, as Negroes and whites, will never be fully accounted for on the environmental hypothesis." But in his personal copy there is a pencilled circle around the mention of Negroes and whites, tied to notes in his handwriting in the margin: "I am less sure of this now (1951)! And still less sure in 1955!—L. M. T."

His religious position was that of an agnostic, but he was not a pessimist with respect to human betterment. Among his papers is a manuscript prepared in 1948, entitled *My Faith*, and intended for publication along with the views of other scientists, though it was not published, presumably because some editor thought it was not sufficiently theistic. He wrote:

At present, our knowledge of human nature and of ways to improve it is not

⁸ Nicholas Pastore, *The Nature-Nurture Controversy*, 1949, 85-95.

keeping pace with progress in the physical sciences. It is for religion, psychology, education, and all the social sciences to set themselves diligently and coöperatively to the task of civilizing man's impulses and emotions, so that nations and peoples everywhere will learn, before it is too late, that it is possible to live together in peace, justice, and good will. . . . From my point of view, the wholehearted dedication of one's life to the promotion of social justice, mutual understanding, racial tolerance, and the equalization of opportunity, is as good evidence of a truly religious spirit as any beliefs one might hold about miracles, immortality, or a personal God.

Thus a farm boy from a most unpromising background endowed with an alert mind, meeting inspiring teachers along the way, made his way to the top of his profession. In the course of his career he broadened the conception of psychology for many who came under his influence, and made it appeal to them as a dignified, exciting, scientifically respectable and socially significant career. This influence lives on in ways that cannot be estimated from the many books and articles that bear Lewis Terman's name as author.

Stanford University

ERNEST R. HILGARD

BOOK REVIEWS

Edited by GEORGE L. KREEZER, Washington University

Perception and the Representative Design of Psychological Experiments. By EGON BRUNSWIK. Berkeley and Los Angeles: Univ. of California Press, 1956. Pp. 146.

Egon Brunswik's last book is an important presentation. A tremendous amount of thought and very little redundancy have gone into most of its 146 pages, and the high (though spotty) conceptual concentration makes for a certain opacity. This, with an astonishingly integrated viewpoint, probably makes casual reading unprofitable.

It is clearly divided into two sections: the methodological first part, *Systematic and representative design of psychological experiments*, has previously appeared in substantially the same form in the *Berkeley Symposium on Mathematical Statistics and Probability* (ed. by Jerzy Neyman, 1949), and as a preprint (1947, 1949); the second part comprises a survey of illustrative functionalist perceptual research entitled *Perception: the ecological generality of its distal aim*.

In Part One, Brunswik develops in detail his distinction between the systematic and representative design in psychological research, illustrated by a compelling series of four prototypical experiments:

(1) The *Galton bar* typifies the structuralist psychophysical judgment of length; statics appears only in error terms, and the main scores yielded are DL's, PSE's, etc.

(2) The "multidimensional" psychophysics of the Gestalt problems is exemplified when the simple Galton bar grows "arrowheads" and becomes the *Mueller-Lyer Illusion*: in the first experiment, physical length and adjacent area had been "artificially tied," so that their relative importance was not determinable; in the second experiment it is revealed that the adjacent area is more powerful in determining S's judgments than was length. This is surprising in that "... it is length which is apt to be singled out by the classical experimenter as the independent variable, both by virtue of its introspective prominence as what may be called the 'figural variable' and by virtue of its being the most conveniently manipulable or measurable aspect of the situation"; *this was concealed by the apparently simple design of the first experiment, which pretends to be univariate but is in reality "clustervariate."* Since the judgment is a "perceptual compromise" between the two "poles of intention" (i.e., length and area), the *constancy ratio* becomes applicable in addition to previous scores (e.g., as a measure of the effect of attitude, etc.).

(3) The *Size Constancy* experiment goes a step further and "unties" the artificial co-variation of the retinal (proximal) stimulus-sizes, and of the objective (distal) stimulus-size: the *standard* is at one distance and the *variable* at another, so that the sizes of the objects, and those of the retinal images they project, are independently variable. The most important score here is the *constancy ratio*, of course. As in experiment (2), our traditional expectations are surprised, and responses are found to be determined more by the *distal stimulus* (for which, ideally, a constancy ratio of 1.0 would be obtained) than by the *proximal stimulus* (ideally, constancy ratio = 0.0). This finding is central to Brunswik's conceptual system for three reasons:

(a) This "distal focussing" (the determination of response by the constant,

stable distal object rather than by its variable and interchangeable proximal projections), is generally characteristic of our perception of things ("thing constancy"), so that we should predict response from the characteristics of the physical object rather than from those of the stimulation of the receptor organs—a revolutionary abandoning of the traditional attempts at internal psychophysics, at establishing chains of causal connections within the organism's physiological structure, a turn from elementaristic structuralist intent to broad, organismic functionalism.

(b) Since constancy ratios in general are *not* usually 1.0 any more than they are 0.0 (and because "ecological validity"—see below—is not perfect), we cannot predict responses unequivocally from the distal stimulus either, but must be content with probabilistic relationships. Indeed, these comprise the very subject matter proper to psychology: stimulus-response probabilities vary not only with individual differences, but with different attitudes, with choice of stimulus and setting employed, etc. Moreover, we cannot hope to isolate the underlying factors because "(in) retrospect, the alleged unidimensionality of the classical experiment . . . must appear fallacious . . . *It is obviously in the nature of things that not all variables except one be held constant* (my italics), if only the scope of experimental observation is extended beyond the immediate conditions . . . (which) must be done if the goal is an inventory of the gross efficiency of organismic functioning . . . rather than a study of the 'fundamental' minutiae of organismic technology" (p. 20).

(c) Because of conceptual errors encouraged by over-generalization from so-called univariate systematic designs, our classical theories are erroneously founded. However, since we cannot simply test *every* conceivable situation, to what variables and in what directions should we broaden our sights? If we could exhaustively study the universe of situations to which the organism can be exposed, we would not be subject to over-generalization. Although this is clearly impossible, we can sample the environment *representatively*, i.e., base our generalization on stimulus situations drawn from the normal ecology of the organism in which we are interested: ". . . representative sampling is extended from the subjects to the objects, from the individuals to the stimulus situations . . . and psychology is conceived of as a fundamentally statistical discipline throughout its entire domain, with 'functional validity' taking its place alongside . . . test validity," and it ". . . becomes possible to ascertain the 'ecological generality' or 'applicability,' of a result along with its counterpart, populational generality or reliability . . . bringing to convergence European academic with Anglo-American statistical tradition" (p. 58). This point, and some of the more formidable Brunswikian terms, may be explicated by our consideration of the last experiment in his heuristic series.

In experiment (4), the "social perception of traits from photographs," not only the subjects, but the stimulus objects as well, are persons. Clearly, we cannot conclude anything usable about the judgment of, say, IQ from photographs, no matter how large and representative the sampling of observers, if the stimuli observed consist solely of American bankers and Australian aborigines, or if the pictures were so selected that the height of the forehead was either held perfectly constant or was forced into perfect correlation with IQ. Representative sampling is thus as necessary for the stimuli as for the observers. The most revealing scores yielded here are, for our present purpose, the eight different uses to which Brunswik puts the *correlation coefficient*: Two are traditional in Differential Psychology, r_{11} and

r_{12} being test (– retest) reliability and test (– other test) validity, respectively. The next three refer to individual factors: $r_{R_1R_2}$, inter-individual reliability (judges' agreement); $r_{RR'}$, intra-individual reliability (judge's self-consistency); $r_{R_A R_B}$, attitudinal validity (the agreement obtained between the response under one set of observational attitudes, and that obtained under another). Most useful, and the one at which we should presumably aim most directly, is *functional validity*, r_{SR} , which measures organismic achievement, *i.e.*, the extent to which an observer correctly responds to the stimulus (here, how well the IQ is judged). Such achievement depends indirectly upon the magnitude of the last two measures: *Ecological reliability*, $r_{SS'}$, measures the constancy of the physically measured stimulus object (how variable is forehead height or IQ from time to time?); *ecological validity*, $r_{S_1S_2}$, measures the extent to which from one stimulus variable we can predict another (such as the correlations between forehead height and IQ, or between the proximity of objects and their mechanical connectedness). The last two are both measures of the physical distal stimulus world alone, but are nevertheless the concern of the psychologist. They could be obtained in complete absence of any observers, were it not essential to maintain a representative sampling of what the organism is likely to confront.

In Part II, Brunswik attempts to fill in for the almost complete absence of suitably representative multivariate perceptual research in the Constancies, and introduces the "perceptual system as an intuitive statistician" in their development; continues his discussion of representative experimental design in terms of some perfectly fascinating hitherto-unpublished factorial studies of the perception of physiognomic qualities (such as *sincerity*, *carelessness*, *likeability*, *intelligence*, etc.) of highly schematized representations of people; produces some very shaky evidence for the ecological validity (and, "therefore," presumably the possible ontogenetically learned basis) of the Gestalt "law" of *proximity*; and (among other things) glances at the problems of ecological representativeness in Clinical Psychology. All of this is done in an incomparable mixture of brilliance, tremendous overgeneralization (of all things!) and a tenacious consistency of viewpoint. A perceptual theory ties this second section together very loosely: organisms are probability-learners, and Perception is "... that relatively primitive, partly autonomous, institutionalized, ratio-morphic [reason-like] subsystem of cognitions which achieves prompt and richly detailed orientation habitually concerning the vitally relevant, mostly distal aspects of the environment on the basis of mutually vicarious, relatively restricted and stereotyped, insufficient evidence in uncertainty-geared interaction and compromise, seemingly following the highest probability for smallness of error at the expense of the highest frequency of precision" (p. 146).

The first part of the book is a mature, polished and magnificently integrated methodology, for which Brunswik claims superior scientific value and a sort of historic inevitability. The second part is relevant only insofar as it gives an indication of the "payoff," of the kinds of research demanded or made possible by the methodology, and there isn't much of such available. In part, this is because: "Representative design in its full scope requires not only a basic theoretical and methodological restructuring but is a formidable task in practice as well. Ideally, it would take concerted research projects of a magnitude hitherto unheard of in experimental psychology" (viii). Since this is admitted to be true, we must approach with caution a position to which such expensive designs are central.

Now, the methodological injunction to use ecologically representative experimental situations has been criticized on other grounds elsewhere; here I would like to point out an inadequacy that I find more important than questions of logic. Most of what were usually considered to be the theoretically important questions in Perception have arisen from the very issues and programs which Brunswik rejects, such as the search for the internal ("merely mediating") physiological variables, for the structuralist elements of experience from which all else we experience could be constructed, from the Gestalt "laws of organization" which proposed to replace the earlier two programs, etc. This is not necessarily a bad "housecleaning"—but, without such motivating questions to guide research, for what reasons should we undertake the endless task of filling out the monstrous tables of ecological correlations and functional validities? "Representative sampling" is a tremendously ingenious methodological concept, but it is *not* an automatic prescription for scientific progress, and in part it conceals a failure of "theoretical nerve," in part a covert theoretical position (this JOURNAL, 67, 1954, 386-388). There is a tendency to accept assertions concerning methodology much more unquestionably and on the basis of far fewer facts and observations than is true of more substantive questions. Perhaps the essential structure of Brunswik's viewpoint and its alternatives will be more apparent if we consider briefly its diametric opposite, the only major *theoretical* alternative which has been introduced in this area since *Gestalttheorie*.

To Gibson, the primary problems of perception remain in essence what they always were—the simple classical psychophysics of space and objects. Ostensibly, Gibson is not a methodologist. No problems of ecologically representative sampling arise, nor is he immersed in a world of probability statistics: except for errors of measurement (usually assumed small), the attributes of the world we see are completely determined by, and in *psychophysical correspondence* with, completely specifiable variables of stimulation. Consequently, whereas Brunswik is essentially an empiricist committed to *probability* learning, reflecting the presumably low probabilities of ecological relationships, Gibson has not really had to be very concerned with learning, as yet; whereas Brunswik requests behemoth multivariate research programs of exhaustive functional surveys, Gibson aims at relatively simple univariate experiments to determine psychophysical correspondences. Where Brunswik saw inherent ambiguity, how can Gibson see certainty? Whatever Gibson's answer, the very existence of an opposition moves one to view Brunswik's "methodological inevitability" with more caution.

The answer lies basically in the very different ways in which they view the stimulus: Brunswik does so in an orthodox manner, and is led to radical experimental design; Gibson employs orthodox experimental design, but a radical treatment of the stimulus. Part of the problem of ambiguity, Gibson solves by (1) the manner in which the stimulus is analyzed, and part by (2) defining his subject-matter so as to avoid areas in which ambiguity seems to be inescapable. Let us see how each of these attempts affects Brunswik's position:

(1) The stimuli to which the organism naturally responds are not the local linear *elements* of the Galton Bar nor the *configuration* of the Mueller-Lyer figure, neither the projected retinal size *nor* the distal object size in the constancy experiments, but rather the total extended sheaf of rays (or retinal image), with all its richness of gradient, motion parallax, etc.; the corresponding percept is not of line or length, but of surface, edge, distance, etc. The stabilized percept is not an "achievement" in

continuously overcoming, say, the changing stimulus of trapezoidal shapes which a rectangular table-top projects from different positions to the moving observer; instead, the constant properties of changing stimulus-gradients (texture-density gradients, etc.) are themselves the simple stimuli for a particular shape-at-a-slant. In Brunswik's terminology, Gibson is saying that if we really examine the stimulation reaching the eye, we will find that both ecological reliability and ecological validity (and for that matter, functional validity as well) are not only high—they are perfect, in principle, at $r_{ss} = r_{s_1s_2} = 1.0$, and he and his colleagues have presented some striking evidence in support of this approach to ecological validity and reliability (although little concerning functional validity). Brunswik, ironically enough, had overgeneralized on the basis of inadequately representative ecological surveys (pictures from magazines, etc.) which were too narrow in scope and sample. The question of *how* good ecological and functional validity really are, is still far from settled, but it is an empirical, inductive, rather than logical, deductive matter.

The fact is, that Brunswik had decided on almost complete ambiguity, and Gibson on complete determinacy, apparently on *a priori* grounds: the study of Perception took much of its initial impetus from the epistemological question of whether or not we can ever know the world with certainty, and both men have strongly concerned themselves with Berkeley's "proof" that we cannot (as a philosophical point, it seems historically very tempting to take an absolutist, all-or-none position on this matter). Of course, this cannot be settled by any factual empirical evidence, despite their efforts to do so (Gibson on occasion even essays the impossible task of completely specifying three dimensions in two, or four in three). As a possible second course tempting Brunswik toward his overgeneralization, we may look to the general positivist orientation with which he was associated. The tendency to achieve maximal communicability through the operational definition has frequently been accompanied by the position that since our scores are expressed as probabilities, the psychological processes themselves must be thought of as probability events. It may be that this viewpoint will eventually prove fruitful in practice; as either an ontological or a procedural dictum, however, it is beyond the province of an empirical science, and without empirical support it can only be viewed as the premature encroachment of a conceptual and methodological orthodoxy.

(2) On the other hand, Brunswik's commitment to ambiguity can prove far less restrictive, in another sense, than Gibson's to determinacy, which tends to lead him to *define* out of the field of perception all responses which are unveridical, of low functional validity (e.g., the optical illusions), or which are ambiguous in the sense that more than one response is likely to occur. Brunswik's position is immune to this weakness of Gibson's. To Brunswik, we can and should study the perception of almost everything. Since ambiguity is omnipresent, there is no reason why we cannot under the same roof study line drawings, reversible figures, persons as stimuli in clinical situations, the "thing constancies," geometric illusions, anything at all at which people look. However, at the same time, he lacks the hope for, or interest in finding the stimuli by which men's responses in these areas are determined, and here we come back to what I feel to be the central weakness in Brunswik's position—an emptiness and lack of prospect, in contrast to the energetic purpose of almost any other "theory" (especially noticeable when contrasted to the optimistic directedness of Gibson's program), an impression of dutiful but unjoyous toil to

fill out the endless tables of organismic functioning. The researches which Brunswik cites or rather, the manner in which he uses them is symptomatic here: they are "demonstrational," presented not for the information yielded but as examples of *how* research should or should not be done.

Is there a purpose or a future which is compatible with Brunswik's position? I think here a comparison of what the two opposed viewpoints actually have in common may prove valuable. First and foremost, although both men were immersed in epistemological philosophy, each has presented good reason for abandoning this discipline as a source of the "theoretically important problems" (e.g., Nature vs. Nurture in Space Perception, etc.): Brunswik, for reasons considered above; Gibson, because he challenges (with signs of success) the traditional arm-chair conclusions about the proximal stimulus which have for centuries always *assumed* low ecological validity; both are equally concerned with the measurement of ecological and functional validity and reliability (although Brunswik tends to ignore the *proximal stimulus*, while Gibson slights *achievement*), and both are equally unimpressed by the promises of physiological speculation as a source of the "important" questions. When we remove these two major traditional high-prestige sources of academic restraint, both tremble on the brink of *utility* (to the arts, commerce, and industry, as well as to other branches of Psychology) as a respectable source of interests. If we set aside their remarkably antipodal philosophically-imposed restraints on subject and method, an equally remarkable area of agreement emerges, and we are left with an infant field of Perception, armed with powerful research tools both old and new, freed from the dead hand of its history, and eager to study consistancies and inconsistancies of perception, wherever, and to whatever extents, they are found.

Cornell University

JULIAN E. HOCHBERG

Shiko Shinrigaku (The Psychology of Thinking III—Thinking in Animals). By TATSURO YATABE. Second edition. Tokyo, Baifukan, 1953. Pp. xl, 464. Supplement by Ryoji Motoyoshi and Koji Sato.

This book is designated as the fourth in a series on the psychology of thinking by Prof. Yatabe. The titles of the first, second, and the third volumes are as follows: *The History of the Psychology of Thinking*, 1948, Pp. xii, 554; *The Psychology of Thinking I—Concepts and Meaning*, 1948, pp. lxxv, 386; and *The Psychology of Thinking II—Relation and Reasoning*, 1949, pp. xxiv, 478.

In *The History of the Psychology of Thinking*, the author tried to grasp the general tendencies in the studies of thinking, especially in the studies of its general principles. It contains the following chapters: 1) Studies of Association, 2) Studies of Memory, 3) Studies of Thought ("Gedankenlehre"), 4) Will and Thinking, and 5) Thinking as Behavior.

In *The Psychology of Thinking I—Concept and Meaning*, and *The Psychology of Thinking II—Relation and Reasoning*, special problems in the thinking of man (adults, children and psychopaths) are discussed. There is a bibliography of 1,090 titles. Chapters include: 1) Formation of Concepts, 2) Understanding of Meaning, 3) Apprehension of Relation ("Relationserfassung"), 4) Formal Reasoning, and 5) Psychology of Problem-Solving. Incidentally, Prof. Yatabe's four volume series is not intended to describe his system of opinions, but to comprise the results

of investigations in the studies of thinking, so that the series can be used as hand-books by investigators in this field. At the present time, there is almost no such work, with its broad point of view and completely detailed descriptions, in any field of psychology.

The first edition of *The Psychology of Thinking III—Thinking in Animals* was published in 1946. In this book, the results of experiments with animals are treated from the point of view of the corresponding problem in the thinking of man, that are treated in "The Psychology of Thinking I and II."

The book is composed of the following five chapters. 1) "Concept in Animals" treats the equivalence-responses in animals. 2) "Apprehension of Relations in Animals" treats the absolute responses and relative responses in animals. 3) "Reasoning in Animals" treats the problems related to the so-called "Maier's Reasoning" and "Krechevsky's Hypotheses." 4) "Thinking and Learning in Animals" treats the problems related to Gestalt theory (K. Duncker, I. Krechevsky), direction theory (N. R. F. Maier), trial-and-error theory (E. Thorndike), conditioned response theory (C. Hull), sign-Gestalt theory (E. Tolman), etc. 5) "Review of Studies made after the War," supplemented in the second edition, treats the results of investigations with animals and children, reported during the period from 1941 to 1953.

The division of the chapters in this book is almost equivalent to that in "The Psychology of Thinking I and II," i.e., "Concept," "Apprehension of Relations (Judgement)" and "Reasoning." This division is originally the category of the thinking process traditionally discussed in logic, and the author himself acknowledges in the preface that it may not be an adequate category in the field of psychology. However, since this category had been used in psychology, and the author's intention in this book is to introduce the psychology of thinking systematically, tracing its historical developments, it is, as the author says in the preface, convenient to make use of this category. Incidentally, in chapter 5, which is supplemented in the 2nd edition by taking the latest developments into consideration, the following classification is given: A) Response Equivalence: 1) Generalization of Situation. 2) Generalization of Principle. 3) Factors in the Generalization of Situation and Principle. B) Variability of Response: 1) Behavior Variability in Concrete Situations. 2) Behavior in the Situation in which Symbolic Function is required. Under these items, the author introduces experimental results, with their problems, discussions and critiques from the standpoint of mechanistic associationism, such as Spence's Stimulus-Response theory, Hull's Generalization theory and Nissen's Approach-Avoidance theory, and from the standpoint of organismic or holistic theories, such as Lashley's Relational theory, Tolman's Sign-Gestalt theory and Bitterman's Configurational theory.

The chief goal of the author is to introduce and discuss the studies in this field systematically, with the greatest fidelity being observed in the description of procedures and of the results of experiments. This fidelity in description should be highly valued. Furthermore, readers who are conducting studies in this field will certainly appreciate the frequently inserted significant discussions about the relationships between the experiments made on men and the experiments made on animals, and the sharp critique of possible fallacies by researchers.

Finally, it should be noted that this book, due to its restricted size and to the

nature of its contents, is not an easy book to read, especially for beginners. Still, this book will be valuable for those who want to know what the present problems in this field are and the problems that have been considered in the past. It is not necessary to read this book from the first page; on the contrary, the reader may begin at any point, with reference to his particular problems. It is to be regretted that this book is not written in a European language.

University of Tokyo

BEN YAGI

New Concepts of Healing: Medical, Psychological, and Religious. By A. GRAHAM IKIN. New York, Association Press, 1956. Pp. xxiii, 262.

This is an unusual book in several respects. In the first place it is a religious book directed to scientists, especially those concerned with humans. At the same time it endeavors to be a scientific book addressed to the clergy, particularly those concerned with divine healing. A book with such a bifurcated task is testimony to the courage of the author. That it does not fail immediately is evidence of her skill. There are other unusual features. The text of the book, 158 pages, was written to an Anglican audience. The 104 pages of appendices, added by Wayne E. Oates, are based exclusively on American activities. The style seems to strive to be semitechnical, but usually varies from popular to metaphysical.

Ikin realizes the welter of material in occultist literature that needs disentangling from the naive evaluations and interpretations often put upon it. For it is exactly this kind of material that has prevented most scientists from taking the subject seriously. She arms herself for the liberation with some of the weapons of science: advocacy of the critical approach, acceptance of the most adequate explanation, respect for the phenomena, and, of course, an eschewal of explaining things away.

On certain questions the author vigorously takes the clergy to task. The Christian ministry must include the welfare of man's body as well as his more eternal interests. The gospel of love in the future should not be substituted for adjustment in the present. The appeal to sacramental help is often an escape from the actual situation. Spiritual healing may be only a sweet resignation to a life of invalidism and being waited upon by others. Authoritarian absolution, such as that practiced in the confessional, lessens moral stature; believing he is forgiven, a person will not have to deal with his basic trouble at all. The value of the book lies in its treatment of the serious perplexities of therapeutic responsibility that confront the clergy. At no time does Ikin substitute the devotional for the intellectual in these considerations.

For the medical and psychological reader, the concepts of healing are not new. The lead is furnished by the work of C. G. Jung, presumably because of his patronage of the Guild of Pastoral Psychology, and a well known chorus line of rather old metaphysicians, Sir James Jeans, C. D. Broad, and Du Noüy supply familiar refrains. Indeed, one of Ikin's principal pleas, if not the *raison d'être* of the book, is not a new concept at all. Close cooperation between healers, spiritual and medical, was part of the Aztec system of medicine before the Spanish invasion. The *Badianus Manuscript* reveals that Aztec medical healers and spiritual healers worked in accord. While the witch doctor chanted to invoke his gods the physician would ply the patient with herbs and potions. The Aztecs, like Ikin, were convinced that doctors were part of the divine source of knowledge.

The American appendices are valuable and interesting, but limited. For instance, the material on Oral Roberts comes solely from two national news magazines. No attempt is made to evaluate his work or his message in terms of medical, psychological, or religious healing. Information on pastoral training in healing does not include those "religious" correspondence schools that offer doctorates, such as the DD, ThD, or PhD, for a few easily completed courses in divine healing. Currently such courses and degrees are causing concern to many established church systems. Surely they are part of the new concept of healing, even if only the bizarre fringe. Nor is the International New Thought Alliance mentioned. Part of this large group consists of the faculty, alumni, and students of the correspondence schools scattered throughout the nation.

The author has illuminated the American scene with a flashlight by centering on questions of academic interest. This reviewer would have preferred a glaring searchlight revealing the fringe as well as the core.

Richmond Professional Institute

DELL LEBO

BOOKS RECENTLY RECEIVED FOR REVIEW

Abramson, H. A., *Neuropharmacology*, New York, Josiah Macy, Jr. Foundation, 1957, pp. 381, \$4.50.

Adair, Aileen, *The Moon Is Full*, New York, Philosophical Library, 1957, pp. 200, \$4.75.

Apostel, L., B. Mandelbrot, and J. Piaget, *Logique et Equilibre*, Paris, Presses Universitaires de France, 1957, pp. 170.

Bates, Marston, P. S. Humphrey, (eds.), *The Darwin Reader*, New York, Charles Scribner's Sons, 1956, pp. ix, 470, \$6.75.

Berg, Charles, *Being Lived by My Life*, London, George Allen & Unwin Ltd., 1957, pp. 256.

Berne, Eric, *A Layman's Guide to Psychiatry and Psychoanalysis*, Second Edition, New York, Simon and Schuster, 1957, pp. xxi, 320.

Beth, W. E., W. Mays, and J. Piaget, *Epistemologie Genetique et Recherche Psychologique*, Paris, Presses Universitaires de France, 1957, pp. 135.

Bier, W. C., S. G. DiMichael, and R. J. McCall, (eds.), *Perception in Present-Day Psychology: A Symposium*, Chicago, American Catholic Psychological Association, 1956, pp. viii, 109, \$1.00.

Bixley, William, *The Guilty and the Innocent*, New York, Philosophical Library, 1957, pp. 176, \$6.00.

Blumel, C. S., *The Riddle of Stuttering*, Danville, The Interstate Publishing Co., 1957, pp. 142.

Bois, J. S., *Exploration In Awareness*, New York, Harper and Brothers, 1957, pp. ix, 212, \$2.75.

Bossard, J. H. S., and E. S. Boll, *One Marriage Two Faiths*, New York, The Ronald Press, 1957, pp. vii, 180, \$3.50.

Bowley, A. H., (with contribution by L. Gardner), *The Young Handicapped Child*, London, E. and S. Livingstone Ltd., 1957, pp. viii, 128, \$3.50.

Brach, Jacques, *Conscience et Connaissance*, Paris, Librairie Marcel Riviere et Cie, 1957, pp. 221.

- Bradfield, J. M., and H. S. Moredock, *Measurement and Evaluation in Education*, New York, The Macmillan Co., 1957, pp. xiv, 509, \$5.50.
- Capa, Cornell, and Maya Pines, *Retarded Children Can Be Helped*, New York, Channel Press, 1957, pp. 159, \$5.00.
- Carmichael, Leonard, *Basic Psychology: A Study of the Modern Healthy Mind*, New York, Random House, 1957, pp. 340, \$3.95.
- Clark, K. E., *America's Psychologists*, Washington, D.C., American Psychology Association, 1957, pp. ix, 247.
- Cleckley, Hervey, *The Caricature of Love*, New York, The Ronald Press, 1957, pp. x, 319, \$6.50.
- Committee on Social Issues, *Psychiatric Aspects of School Desegregation*, New York, Group for the Advancement of Psychiatry, 1957, pp. 95, \$1.00.
- Cronbach, L. J., and G. C. Gleser, *Psychological Tests and Personnel Decisions*, Urbana, University of Illinois Press, 1957, pp. xii, 165, \$3.50.
- Cumming, Elaine, and J. Cumming, *Closed Ranks: An Experiment in Mental Health Education*, Cambridge, Harvard University Press, 1957, pp. xx, 192, \$3.50.
- Cutts, N. E., Nicholas Mosely, *Teaching the Bright and Gifted*, Englewood Cliffs, New Jersey, Prentice-Hall, 1957, pp. xiv, 268.
- Davidson, Donald, and Patrick Suppes (in collaboration with Sidney Siegel), *Decision Making: An Experimental Approach*, Stanford, California, Stanford University Press, 1957, pp. 121, \$3.25.
- Davidson, H. A., *Guide to Medical Writing*, New York, The Ronald Press, 1957, pp. vii, 338, \$5.00.
- DeHann, R. F., and R. J. Havighurst, *Educating Gifted Children*, Chicago, The University of Chicago Press, 1957, pp. ix, 276, \$5.00.
- D'Evelyn, K. E., *Meeting Children's Emotional Needs*, Englewood Cliffs, New Jersey, Prentice-Hall, 1957, pp. x, 176.
- Eng, Helga, *The Psychology of Child and Youth Drawing*, New York, The Humanities Press, 1957, pp. vii, 205, \$9.50.
- Evans-Wentz, W. Y., *Tibetan Book of the Dead*, Third Edition, New York, Oxford University Press, 1957, pp. lxxxiv, 249, \$4.00.
- Ewalt, J. R., E. A. Strecker, and F. G. Ebaugh, *Practical Clinical Psychiatry*, New York, McGraw-Hill Book Company, 1957, pp. xiv, 457, \$8.00.
- Eysenck, H. J., *Sense and Nonsense in Psychology*, Baltimore, Penguin Books, 1957, pp. 349, \$.85.
- Fairfield, Letitia, *Epilepsy*, New York, Philosophical Library, 1957, pp. 159, \$4.75.
- Ferguson, T., and J. Cunnison, *In Their Early Twenties*, London, Oxford University Press, 1957, pp. vi, 110, \$2.00.
- Foster, J. H., *Sex-Variant Women in Literature*, New York, Vantage Press, 1957, pp. 412.
- Foulkes, S. H., and E. J. Anthony, *Group Psychotherapy*, Baltimore, Penguin Books, 1957, pp. 263, \$.85.
- Funkenstein, D. H., S. H. King, and M. E. Drolette, *Mastery of Stress*, Cambridge, Harvard University Press, 1957, pp. xv, 329, \$6.00.
- Garfield, Sol L., *Introductory Clinical Psychology*, New York, The Macmillan Company, 1957, pp. xiii, 469, \$6.00.

- Gesell, Arnold, F. L. Ilg, and L. B. Ames, *Youth: The Years from Ten to Sixteen*, New York, Harper and Brothers, 1956, pp. xv, 542.
- Greenberg, J. H., *Essays in Linguistics*, Chicago, University of Chicago Press, 1957, pp. vii, 108, \$3.00.
- Greenwood, David, *Truth and Meaning*, New York, Philosophical Library, 1957, pp. xii, 114, \$3.75.
- Guest, Lester, *Beginning Statistics*, New York, Thomas Y. Crowell Company, 1957, pp. ix, 255, \$4.00.
- Harris, D. B., (ed.), *The Concept of Development*, Minneapolis, University of Minnesota Press, 1957, pp. x, 287, \$4.75.
- Hart, Bernard, *The Psychology of Insanity*, Fifth Edition, New York, Cambridge University Press, 1957, pp. xi, 127, \$1.75.
- Hartley, E. L., and R. E. Hartley, *Outside Readings in Psychology*, Second Edition, New York, Thomas Y. Crowell Company, 1957, pp. ix, 499, \$2.50.
- Heath, Spencer (with foreword by John Chamberlain), *Citadel, Market and Altar*, Elkridge, Maryland, The Science of Society Foundation, 1957, pp. xix, 259, \$6.00.
- Henrysson, Sten, *Applicability of Factor Analysis in the Behavioral Sciences*, Stockholm, Almqvist & Wiksell, 1957, pp. 156.
- Hewitt, D. W., *Alcoholism*, Philadelphia, Lea & Febiger, 1957, pp. 112.
- Hilgard, Ernest, *Introduction to Psychology*, Second Edition, New York, Harcourt, Brace and Company, 1957, pp. xv, 653.
- Himwich, Harold, *Alcoholism: Basic Aspects and Treatment*, Washington, D.C., The American Association for the Advancement of Science, 1957, pp. viii, 220, \$5.75.
- Himwich, Harold, *Tranquilizing Drugs*, Washington, D.C., The American Association for the Advancement of Science, 1957, pp. viii, 205, \$5.00.
- Hutt, M. L., and R. G. Gibby, *Patterns of Abnormal Behavior*, Boston, Allyn and Bacon, 1957, pp. xvi, 452, \$6.00.
- Jersild, A. T., *The Psychology of Adolescence*, New York, The Macmillan Company, 1957, pp. xii, 438, \$5.00.
- Jones, M. H., *Executive Decision Making*, Homewood, Illinois, Richard D. Irwin, 1957, pp. xvii, 485, \$6.00.
- Kahn R. L., and Charles Cannell, *The Dynamics of Interviewing*, New York, John Wiley & Sons, 1957, pp. x, 368, \$7.75.
- Katz, Barney, *You Can Have a Better Marriage*, New York, The American Press, 1956, pp. 177, \$3.00.
- Lafite, Paul, *The Person in Psychology: Reality or Abstraction*, New York, Philosophical Library, 1957, pp. x, 233, \$6.00.
- Langer, Marion, *Learning to Live as a Widow*, New York, Julian Messner, 1957, pp. 255.
- Linschoten, J., *Strukturanalyse der Binokularen Tiefenwahrnehmung*, Groningen, Holland, J. B. Wolters (New York, Gregory Lounz), 1957, pp. xix, 573, \$10.00.
- Lonergan, B. J. F., *Insight: A Study of Human Understanding*, New York, Philosophical Library, 1957, pp. xxx, 785, \$10.00.
- Louttit, C. M., (ed.), *Clinical Psychology of Exceptional Children*, New York, Harper and Brothers, 1957, pp. xii, 573, \$6.00.
- Lumeij, J. L. J., *The Methods of Psychology and Psychiatry*, Netherlands, Van Gorcum & Company, 1957, pp. 243.

- Meili-Dworetzki, Gertrud, *Das Bild des Menschen in der Vorstellung und Darstellung des Kleinkindes*, Bern, Verlag Hans Huber, 1957, pp. 136.
- Midelfort, C. F., *The Family in Psychotherapy*, New York, McGraw-Hill Book Company, 1957, pp. ix, 203, \$6.50.
- Mounier, Emmanuel, *The Character of Man*, New York, Harper and Brothers, 1957, pp. x, 341, \$6.00.
- Murtagh, J. M., and Sara Harris, *Cast the First Stone*, New York, McGraw-Hill Book Company, 1957, pp. 307, \$4.50.
- Neustatter, W. L., *The Mind of the Murderer*, New York, Philosophical Library, 1957, pp. viii, 232, \$6.00.
- Norris, Miriam, P. J. Spaulding, and F. H. Brodie, *Blindness in Children*, Chicago, The University of Chicago Press, 1957, pp. xv, 173, \$3.00.
- Packard, Vance, *The Hidden Persuaders*, New York, David McKay Company, 1957, pp. vii, 275, \$4.00.
- Pavlov, I. P., *Experimental Psychology and Other Essays*, New York, Philosophical Library, 1957, pp. 653, \$7.50.
- Phillips, William, (ed.), *Art and Psychoanalysis*, New York, Criterion Books, 1957, pp. xxiv, 552, \$8.50.
- Pressey, S. L., and R. G. Kuhlen, *Psychological Development Through the Life Span*, New York, Harper and Brothers, 1957, pp. xxiii, 654.
- Rickett, Allyn, and Adele Rickett, *Prisoners of Liberation*, New York, Cameron Associates, 1957, pp. xiii, 15-288, \$4.75.
- Rieker, Hans-Ulrich, (transl. by A. J. Pomerans), *The Secret of Meditation*, New York, The Philosophical Library, 1957, pp. 176, \$6.00.
- Riverso, Emmanuele, *Metafisica e Scientismo: Con Un'Appendice Sulla Logica di C. S. Peirce*, Napoli, Istituto Editoriale Del Mezzogiorno, 1957, pp. 149.
- Robaye, Francine, *Niveaux D'Aspiration et D'Expectation*, Paris, Presses Universitaires de France, 1957, pp. 223.
- Rohde, A. R., *The Sentence Completion Method: Its Diagnostic and Clinical Application to Mental Disorders*, New York, The Ronald Press, 1957, pp. xii, 301, \$7.50.
- Ronchi, Vasco, (transl. and revised by Edward Rosen), *Optics: The Science of Vision*, New York, New York University Press, 1957, pp. ix, 360, \$10.00.
- Rosen, Elizabeth, *Dance in Psychotherapy*, New York, Bureau of Publications, Columbia University, 1957, pp. xx, 178, \$4.50.
- Ross, Ralph, and Ernest Van Den Haag, *The Fabric of Society*, New York, Harcourt, Brace and Company, 1957, pp. xv, 777.
- Sachs, Wulf, *Black Anger*, New York, Grove Press, 1957, pp. 324, \$1.75.
- Sears, Robert, E. E. Maccoby, and Harry Levin, *Patterns of Child Rearing*, New York, Row, Peterson and Company, 1957, pp. vii, 549.
- Shneidman, E. S., and N. L. Farberow, (eds.), *Clues to Suicide*, New York, McGraw-Hill Book Company, 1957, pp. xii, 227, \$5.50.
- Simon, H. A., *Models of Man*, New York, John Wiley and Sons, 1957, pp. xiv, 287, \$5.00.
- Spinoza, Baruch, (ed. by D. D. Runes), *The Road to Inner Freedom*, New York, Philosophical Library, 1957, pp. 215, \$3.00.

- Steckle, L. C., *Problems of Human Adjustment*, New York, Harper and Brothers, 1957, pp. xi, 419.
- Stone, Joseph, and Joseph Church, *Childhood and Adolescence*, New York, Random House, 1957, pp. xvii, 456, \$6.50.
- Stott, D. H., *Unsettled Children and Their Families*, New York, Philosophical Library, 1957, pp. 240, \$6.00.
- Super, D. E., *The Psychology of Careers*, New York, Harper and Brothers, 1957, pp. x, 362, \$5.75.
- Szasz, T. S., *Pain and Pleasure*, New York, Basic Books, 1957, pp. xvi, 301, \$5.50.
- Tashman, H. F., *Today's Neurotic Family*, New York, New York University Press, 1957, pp. vii, 214, \$3.95.
- Taylor, R. W., (ed.), *Life, Language, and Law*, Ohio, Antioch Press, 1957, pp. xii, 223, \$4.50.
- Templin, M. C., *Certain Language Skills in Children*, Minneapolis, The University of Minnesota Press, 1957, pp. xviii, 183, \$4.00.
- Tomkins, S. S., and J. B. Miner, *The Tomkins-Horn Picture Arrangement Test*, New York, Springer Publishing Company, 1957, pp. xv, 383, \$10.00.
- Travis, L. E., (ed.), *Handbook of Speech Pathology*, New York, Appleton-Century-Crofts, 1957, pp. viii, 1088.
- Traxler, A. E., *Techniques of Guidance*, Rev. Ed., New York, Harper and Brothers, 1957, pp. xiii, 374, \$6.00.
- Vincent, C. E., *Readings in Marriage Counseling*, New York, Thomas Y. Crowell Company, 1957, pp. xii, 500, \$4.50.
- Walker, Kenneth, *Patients and Doctors*, Baltimore, Penguin Books, 1957, pp. 182, \$.85.
- Wallerstein, R. S., *Hospital Treatment of Alcoholism*, New York, Basic Books, 1957, pp. x, 212, \$5.00.
- Wallin, J. E. W., *Mental Deficiency in Relation to Problems of Genesis, Social and Occupational Consequences, Utilization, Control, and Prevention*, Brandon, Vermont, Journal of Clinical Psychology, 1956, pp. 188.
- Walsh, H. G. (with chapters by John Marshall), *Physiology of the Nervous System*, New York, Longmans, Green and Co., 1957, pp. xv, 563.
- Warner, S. J., *The Urge to Mass Destruction*, New York, Grune & Stratton, 1957, pp. xii, 188, \$3.50.
- Williams, R. J., *Biochemical Individuality*, New York, John Wiley & Sons, 1956, pp. xiii, 214.
- Woodbury, Angus, *Comfort for Survival*, New York, Vantage Press, 1957, pp. 104, \$3.00.
- Wormhoudt, Arthur, *Hamlet's Mouse Trap*, New York, Philosophical Library, 1957, pp. 221, \$3.50.
- Yates, D. H., *Psychology You Can Use*, New York, Thomas Y. Crowell Co., 1957, pp. vii, 248, \$3.95.
- Zaehner, R. C., *Mysticism Sacred and Profane*, New York, Oxford University Press, 1957, pp. xviii, 256, \$6.75.
- Zander, Alvin, A. R. Cohen, and Ezra Stotland, *Role Relations in the Mental Health Professions*, Ann Arbor, Michigan, Instit. for Soc. Research, 1957, pp. vii, 211.



THE AMERICAN JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXX

DECEMBER, 1957

No. 4

THE EFFECT OF RETINAL AND PHENOMENAL ORIENTATION ON THE PERCEPTION OF FORM

By IRVIN ROCK and WALTER HEIMER, New School for Social Research

The relative importance of retinal and environmental orientation in form perception was studied in a previous experiment.¹ Ordinarily a change in orientation gives rise to a phenomenal change, but with the head upright there is a change in orientation of the retinal image as well as in orientation of the form in relation to the environmental upright. In experiments in which these two factors were isolated from each other either by tilting *O*'s head or by tilting the scene by means of a prism, the retinally normal orientation no longer coincides with the environmentally normal (upright) one. In spite of contrary indications from observations in daily life and from a few previous experiments, the results showed decisively that changing a form's orientation from its normally upright environmental position (with retinal orientation remaining normal) makes for greater difficulty in recognition than changing its orientation on the retina (with environmental orientation remaining normal).

In the present paper we report the results of recent experiments which show that this finding is not limited to the use of the particular technique previously employed, and then we turn to the problem raised by the difference between our results and other data which seem to show that it is primarily the retinal factor which affects the perception of form.

EXPERIMENTS I-III: RECOGNITION OF SINGLE FIGURES

In one of the techniques developed to study this problem, *O* had to select one of several alternatives that included two replicas of the original figure (critical figures). One was phenomenally upright either because

* Received for publication April 27, 1956.

¹ Irvin Rock, The orientation of forms on the retina and in the environment, this JOURNAL, 69, 1956, 513-528.

it remained objectively upright or because it was upright in relation to a tilted scene; the other was retinally normal because it was always tilted by the same amount as the head of the *O*. The *O* tended to select the phenomenally upright figure. In view of the surprising nature of these results, the question arises whether they are merely an artifact of the particular procedure employed. Presumably *O* perceives both of the two critical figures and selects the one that obtrudes itself on him, so to speak, as similar to the previously seen figure. It might be argued that *O*'s choice reflects not so much the change in appearance of one of the two critical figures but rather which of the two figures, both correctly perceived, *O* thinks is the right answer. The right answer would be held to be the figure that has the same objective orientation as the original, namely, upright.

This argument is not very plausible because, for one thing, if it were true we would expect individual *O*s consistently to choose the upright alternative in all three test-cards and this is rarely the case. Secondly, most *O*s are not at all aware that there are two alternatives in each test-card. To allay any doubt on this point, we varied the procedure in such a way as to eliminate this possible criticism. A technique was employed, much like the one used by Oetjen,² that required *O* to indicate from a tilted position whether figures looked like those seen previously from an upright position.

Instead of selecting one figure from among several on a given test-card, *O* was shown a series of 18 test-figures one at a time. The series consisted of the three sets of test-figures used in the previous experiments.³ There were, therefore, 12 non-critical figures and 6 critical figures—2 for each training-figure, 1 upright and the other tilted 90° to the same side as *O*.

O was instructed to report whether every figure was recognized as one of the three shown previously. The series was shown in a prearranged order in which the critical figures appeared in the third, sixth, eighth, eleventh, fifteenth, and seventeenth positions. At each of these points, the critical figures were exposed in two positions. For half of the *O*s, the positions were: A upright, B tilted, C upright, A tilted, B upright, C tilted; and for the other half the positions were reversed.

During initial training, the exposure-times, controlled by means of a tachistoscope, were 1 sec. for the three training-figures and either 0.1 sec. or 0.05 sec. for the test-figures. In the earlier study, longer training- and test-intervals were used, but those intervals would have led to the recognition of every test-figure if used here. For similar reasons, *O* was not told during the training-exposures that subsequently his memory for these figures would be tested.

Experiment I. Experiment I was a control in which both retinal and environmental

² Friedrich Oetjen, Die Bedeutung der Orientierung des Lesestoffs für das Lesen und der Orientierung von sinnlosen Formen für das Wiedererkennen derselben, *Z. Psychol.*, 71, 1915, 321-355.

³ Rock, *op. cit.*, see Fig. 3, p. 520.

disorientation were used. It provided data against which to measure the separate effects. The experiment was done with *O* remaining upright during training and test trials. One version of each critical test-figure was presented upright and the other was tilted 90°. There were two conditions. In Condition I, one version of each critical figure was tilted to the right and in Condition II one version of each was tilted to the left. Fifteen *O*s were employed in each condition with a 0.1-sec. test-exposure and an additional 10 *O*s in each condition with a 0.05-sec. test-exposure.

Results. Table I gives the number of recognitions of the critical figures when presented upright (U) and when presented tilted (T). Critical Figure C did not yield a sufficiently clear-cut difference in the recognizability of its upright and

TABLE I
NUMBER OF RECOGNITIONS OF CRITICAL FIGURES IN EXPERIMENT I
(Head upright; figures upright or tilted both retinally and environmentally)

Condition	Figure					
	A		B		both	
	U	T	U	T	U	T
I. Figures tilted right						
1/10 sec. exposure ($N=15$)	11	6	10	4	21	10
1/20 sec. exposure ($N=10$)	7	3	7	3	14	6
II. Figures tilted left						
1/10 sec. exposure ($N=15$)	13	5	10	5	23	10
1/20 sec. exposure ($N=10$)	8	3	6	3	14	6
Total	39	17	33	15	72	32
Percentage	78	34	66	30	72	32

tilted positions, hence the results with it are not reported. (For some reason this figure is almost as easy to recognize when tilted as when upright.)

The total number of possible recognitions of each critical figure for Conditions I and II combined is 50. The percentages in the bottom row are computed from this total. As can be seen, upright critical figures were recognized 72% of the time and tilted ones 32% of the time. This difference is significant at the 1% level. This result affords a measure of the disturbance to recognition, under the conditions of this experiment, brought about by a 90° tilt of these two figures.

The question now to be answered is: To what extent is this disturbance caused by the fact that the tilted figure gives rise to an image disoriented on the retina and to what extent is it caused by the fact that the tilted figure is disoriented in the environment? The answer to this question was sought in the next experiment.

Experiment II. The procedure in Experiment II was identical in every respect with that of Experiment I except that now after viewing the figures in training, *O* tilted his head 90° and viewed the test-figures with his head lying on its side on a table. There were two conditions. In Condition I, *O*'s head was tilted 90° to the right and in Condition II it was tilted 90° to the left. There were 15 *O*s employed in each condition with a 0.1-sec. test-exposure and an additional 10 *O*s in each condition with a 0.05-sec. test-exposure.

Results. Table II gives the number of recognitions of the critical figures when

presented upright in the environment (U) and when tilted 90° so as to be normal in retinal orientation (RN). The total number of possible recognitions of each critical figure for Conditions I and II combined is 50. The percentages in the bottom row are computed from this total.

As can be seen, upright figures were recognized 69% of the time and tilted figures were recognized 42% of the time. Again the environmentally upright

TABLE II
NUMBER OF RECOGNITIONS OF CRITICAL FIGURES IN EXPERIMENT II
(Head tilted 90° on test)

Condition	Figure					
	A		B		both	
	U	RN	U	RN	U	RN
I. Head right						
1/10 sec. exposure ($N=15$)	12	8	9	6	21	14
1/20 sec. exposure ($N=10$)	8	4	5	3	13	7
II. Head left						
1/10 sec. exposure ($N=15$)	11	9	11	4	22	13
1/20 sec. exposure ($N=10$)	5	3	8	5	13	8
Total	36	24	33	18	69	42
Percentage	72	48	66	36	69	42

figures are favored. Chi-squared, comparing the distributions of success and failure of the two modes of presentation of critical test-figures, is significant at the 1% level.

These results differ only slightly from the control results of Experiment I. Upright figures, which are now retinally tilted, lose only 3% in comparison with the upright ones that are retinally normal (72% vs. 69%). Similarly, tilted figures which now are retinally normal gain only 10% with respect to those in Experiment I that are retinally tilted (32% vs. 42%). Both comparisons suggest that the retinal factor is not playing much of a role. The slight differences in percentages are not significant.

The results of the two experiments taken together show that when a figure remains upright in the environment, there is no appreciable loss at all caused by the retinal tilt of 90° . On the other hand, when a figure is tilted environmentally by 90° but remains retinally normal, there is a significant drop of 30% in the recognition-score (42% vs. 72%).

Experiment III. Experiment III employs a procedure that is the converse of that used in Experiment II. *O* viewed the training-figures from a 90° tilted position and the test-figures from an upright position.

Results. The results for a total of 65 *O*s for head left and head right conditions, both figures combined, were: environmentally upright critical figures were recognized 68% of the time and retinally normal critical figures recognized 53% of the time. This difference is significant at close to the 1% level. Again there is no loss in recognition of upright figures despite the change in the retinal orientation. The 68% correct in this experiment is to be compared with 72% in the control experiment. There is, however, a significant gain, over the result of Experiment I, in the

recognition of tilted figures (53% vs. 32%). The gain as a result of similar retinal orientation is greater in this experiment than it was in Experiment II. Except for this fact the results of this procedure confirm those of Experiment II.

Thus the results of Experiments I, II, and III show by an entirely different technique that environmental change of orientation is quite disturbing to recognition whereas retinal change is either without effect or plays only a minor role. It is hard to see how these results can be regarded as due to any special feature of the procedure which might limit their significance. Either *O** recognized a critical figure or he did not. If he recognized a tilted figure, he reported it although he might add that it was differently oriented. Such reports were scored as successes. We must assume, therefore, that the absence of a response to a tilted figure represents a genuine failure of recognition. Since the results are similar to those of the earlier experiments, in which a choice of figures was offered to *O*, the earlier results may now be regarded as authentic. Indeed, in the light of both sets of results, the same conclusion may be reached in regard to the results of the experiments with ambiguous figures. The three different procedures give confirmatory results.

EXPERIMENTS IV-VI: RECOGNITION OF COMPLEX FIGURES

We turn now to the problem raised by the difference between our results showing a dominant role of the environmental factor and the results of demonstrations, such as those obtained by Köhler, showing that the retinal factor seems to play the dominant role.⁴ One obvious difference between these procedures and ours is the kind of figure used. A word in script or a human face is much more complex than any of the figures we employed. More complex figures are presumably more difficult to recognize, and, being more difficult to recognize, they would be more adversely affected by retinal disorientation. If a simple form is used, such as single printed letter of the alphabet, it is easily recognized, even when retinally disoriented. Our results with simple figures bear out this conclusion.

In an effort to investigate the part played by difficulty in recognition, a series of experiments was undertaken using fragmented figures of the kind devised by Street.⁵ These figures have been made difficult to recognize even when they are upright, hence they are particularly well suited for study of this problem.

Experiment IV. Experiment IV provided control data against which to measure

⁴ Wolfgang Köhler, *Dynamics in Psychology*, 1940, 15-30.

⁵ R. F. Street, *A Gestalt Completion Test*, 1931, 65.

the separate effects of the two modes of disorientation. *O* remained upright and viewed the figures which were either upright or tilted 90°, as in Experiment I. A training exposure was not necessary because the figures represent familiar objects.

On the basis of pre-testing, eight figures were selected which were of such difficulty that they would be recognized occasionally when tilted, but not all the time in both positions. They were divided into two sets of four each. In Set A were the figures of a car and typewriter used by Leeper, a ship used by Verville and Cameron, and a piano designed by John Urban. In Set B were Street's figures of the boy on bicycle, a horse and carriage, an elephant used by Leeper and a bull designed by Urban.⁶ One group of 20 *O*s was shown all eight figures upright. Another group of 10 *O*s was shown the figures of Set A tilted 90° to the left; a third group of 10 *O*s was shown the figures of Set B tilted 90° to the left. (These two groups of 10 *O*s were shown the four figures comprising the other set as well, as a separate experiment.) The order of presentation of the figures was randomized for each *O*.

O was first shown an easy sample in order to explain what is involved in

TABLE III
NUMBER OF RECOGNITIONS OF COMPLEX FIGURES IN EXPERIMENT IV
(Head upright)

Condition	Set A				Set B				All	%
	car	type- writer	ship	piano	bull	ele- phant	boy on bicycle	horse & car- riage		
Figures upright (<i>N</i> =20)	13	6	14	20	18	12	8	15	106	66
Figures tilted (<i>N</i> =10)	1	1	0	2	2	1	1	2	10	12

recognizing such figures. Each figure was then presented for a maximum of 20 sec. No help was given and *O* was scored either right or wrong for each figure.

Results. The results are given in Table III. Of the upright figures, 66% (106 out of 160) were correctly reported, while of the tilted figures only 12% (10 out of 80) were seen. This difference is highly significant.

Experiment V. In Experiment V, *O*'s head was tilted 90° to the right. Thus in both this experiment and in Experiment IV, there was 90° counter-clockwise retinal disorientation. Except for this difference, the procedure was identical with that used in Experiment IV. The *O*s were divided into two groups of 20 each. One group

⁶ Robert Leeper, A study of a neglected portion of the field of learning—the development of sensory organization, *J. genet. Psychol.*, 46, 1935, 41-75; Elinor Verville and Norman Cameron, Age and sex differences in the perception of incomplete pictures by adults, *J. genet. Psychol.*, 68, 1946, 149-157.

was shown the figures of Set A upright and the figures of Set B tilted 90° ; the other group was shown the figures of Set B upright and those of Set A tilted 90° . For half the *O*s, Set A was shown first and for half, Set B was shown first. The order of presentation of figures within each set was randomized for each *O*.

Results. In Table IV the possible score for a given figure when upright or tilted is 20. Altogether, figures upright in the environment, and therefore retinally tilted, were recognized 51% of the time (82 out of a possible 160), and figures tilted 90° ,

TABLE IV
NUMBER OF RECOGNITIONS OF COMPLEX FIGURES IN EXPERIMENT V
(Head tilted 90°)

Condition	Set A				Set B				All	%
	car	type- writer	ship	piano	bull	ele- phant	boy on bicycle	horse & car- riage		
Figures upright (<i>N</i> =20)	14	6	7	15	18	6	3	13	82	51
Figures tilted (<i>N</i> =20)	9	4	7	13	11	2	4	11	61	37

and therefore retinally normal, were recognized 37% of the time (61 out of 160). Chi-square comparing the two distributions of success and failure is significant beyond the 2% level.

From a recognition score of 66% obtained in Experiment IV when both *O* and figures were upright, there is a drop to 51% when *O* is tilted and the figures remain upright. This difference of 15% yields a Chi-square significant at the 1% level. When *O* and figures are both tilted 90° , maintaining retinal normalcy, 37% are correct; and finally, when only the figures are tilted producing *both* environmental and retinal change, the percentage drops to 12. In this latter comparison retinal normalcy produces a difference of 25%, significant at the 1% level. Both comparisons, in contrast to the results of Experiments IX and X with nonsense-figures, clearly show that for such difficult figures retinal disorientation definitely interferes with recognition. To this extent, therefore, our belief is confirmed that we failed, on the whole, to obtain such results in the previous experiments because relatively simple figures were used.

Another difference between our results and the observations of other experimenters is that we found orientation in the environment to play a very considerable role in form perception whereas we do not experience the slightest trouble reading text which is upside down or tilted 90° in the environment (provided it remains retinally normal). The difficulty of the material, as exhibited in the last series, cannot account for the difference. From the level of 66% obtained in Experiment IV, recognition drops to 37% in Experiment V when the figures are tilted environmentally but remain normal retinally. This difference is highly significant and it is a

greater difference than that produced by retinal disorientation alone. How can we explain this contradiction?

Suppose that *O* with his head tilted sees an unfamiliar figure in a familiar setting. Prior to recognition, he does not know what it is he is confronted with. It is likely that, although he is tilted, he will organize the pattern in such a way that its top agrees with the direction which is up in the environment. This is plausible because with no information to the contrary there is no reason to see an as yet unfamiliar object as tilted in the environment. This selection of an orientation in turn influences the form's appearance. If the figure is in fact upright the directions are selected correctly whether *O* is himself upright or not. If it is tilted, the directions are, of course, incorrectly chosen and this results in a changed phenomenal shape. In the case of Street's type of figure, the same argument will hold. On the other hand, if the *O* knows that the figure is tilted, the choice of an incorrect orientation might be avoided, and as a result recognition might not be seriously disturbed. With a very familiar shape, however, such as a face or a word in text or script, one may

TABLE V
NUMBER OF RECOGNITIONS OF COMPLEX FIGURES IN EXPERIMENT VI
(Information given prior to test)

Condition	Set A				Set B				All	%
	car	type- writer	ship	piano	bull	ele- phant	boy on bicycle	horse & car- riage		
Figures upright (<i>N</i> =20)	16	6	6	18	17	11	3	14	91	57
Figures tilted (<i>N</i> =20)	17	11	15	20	17	8	8	15	111	69

recognize at the outset the general class of object and, as a result, one does not make a mistake in the choice of direction. This, we believe, may explain why with such shapes environmental disorientation alone does not have a disturbing influence.

To test this hypothesis, an experiment was undertaken in which *O* was informed about the orientation of the figures. It was expected that such information could serve to overcome the difficulty in recognition of environmentally tilted figures by enabling *O* to organize the figures in reference to the correct directions.

Experiment VI. The procedure of Experiment VI was identical to that used in Experiment V with the single exception that now *O* was told prior to being shown the four figures of a given set either that "the following figures will all be upright" or that "the following figures will all be tilted 90° to the same side as your head." Thus in every case *O* knew which side of the figure was the top. As in Experiment V, *O*'s head was tilted 90° to the right throughout. Twenty *O*s were shown Set A figures upright and Set B figures tilted and 20 were shown Set A figures tilted and Set B figures upright.

Results. The results in Table V show that the figures, when upright with *O* explicitly told that they were upright, were recognized 57% of the time (91 out of

a possible 160). This is not greatly different from the score of 51% obtained for the upright figures in Experiment V without such explicitly given information. We should not expect any improvement for *upright* figures if it is true that with no information to the contrary (as in Experiment V) *O* will tend to see a form in such a way that its top is the side which is 'up' in the environment.

When the figures were tilted, with *O* explicitly told that they were tilted, they were recognized 69% of the time (111 out of a possible 160). Thus the large decrement in recognition in the parallel condition in Experiment V, without information, is now completely eradicated. These *O*s do as well as is to be expected under optimal control conditions when both *O* and figures are upright (66%). The slightly superior score of 69% along with the difference noted above (57% vs. 51%) suggests that the *O*s in this experiment may have been a little better, on the whole, in recognizing the type of figure used.

In other words, the information completely eliminates the disturbance created by environmental tilt in Experiment V. (Exactly what process occurs which makes it possible for this information to lead to a changed phenomenal orientation of the figure is not understood in this case any more than elsewhere in perception where a special attitude or set leads to sensory reorganization.) Since the figures in this condition are retinally normal, there is now no disturbance at all. This result, therefore, confirms our hypothesis. It may be noted that the upright figures which are recognized 57% of the time are recognized significantly less frequently than the tilted ones. Chi-squared for the differences between the two distributions is significant at the 2% level. Since information is provided in both cases as to which direction of the figures is the 'top,' this difference can only mean that retinal disorientation of the upright figures is producing the disturbance—confirmation that, with figures as difficult as these, retinal disorientation interferes with recognition.⁷

It now seems clear that the orientation of a form in the environment influences its appearance because the directions in the environment serve as a rule to define the directions of the form. The important factor, then, which should be contrasted with the retinal factor, is not environmental orientation as such but the selection of the parts of a form which are to be its phenomenal top, bottom and sides. With no information or set to the

⁷ In the previous paper an experiment by Oetjen (*op. cit.*, 321-355) was described in which both text and unfamiliar nonsense-figures were used in a procedure similar to that of Experiments I and II. He found in both cases that forms were easily recognized as long as they were retinally normal, whether they were tilted in the environment or not; but were difficult to recognize when retinally tilted. We have dealt with the use of text, but his results with nonsense-figures remain unexplained. Since he used a long series of fairly complex figures, it is not surprising that retinal disorientation proved disturbing. It is not at all clear, however, why he did not find that disorientation of these forms in the environment played much of a role. It is possible that his *O*s became sophisticated about the fact that the forms were shown in varying orientations, since each *O* served over a period of 12 days. With unlimited time for inspection of the test-figures, *O* could then search for the 'top' of each. The resulting selection of the correct orientations would, of course, eliminate the disturbance of environmental disorientation as in our Experiment VI.

contrary, environmental directions are applied; but knowledge to the contrary can alter this whether it is given by instruction or by familiarity with the type of material employed. There is still another way in which the phenomenal directions of a form may be selected which will be brought out in the following discussion.

EXPERIMENTS VII AND VIII: RECOGNITION OF FIGURES ON A HORIZONTAL SURFACE

When the earlier experiments were being conducted,⁸ in which *O* selected the critical figure that appeared most similar to the training-figure, a technique was sought which would provide a measure of the role of retinal orientation with orientation in the environment held constant. The following experiment was the result.

Experiment VII. The general procedure of the earlier experiments (III-VI) was employed with the following changes. After *O* inspected the three training-figures he was required to view from a standing position the test-cards which rested on the floor. *O* bent his head over a circular cardboard containing a small aperture in the center, so as to look directly down at the test-figures. One critical figure was retinally normal and the other figure was turned 90° . Since the cards were in a horizontal plane, neither critical figure was favored as far as environmental orientation is concerned because the figures' orientation to the upright of the environment does not change when it is rotated in this plane.⁹ Other possible frames of reference were also excluded, *e.g.* the figures were so exposed that they could only be seen on a circular field through the aperture.

Results. The retinally normal figure was strongly favored over the tilted one. For approximately 25 *O*s, 89% of the selections were of this normal alternative. This result raises the following problem. If the assumption were correct that in this experiment turning the figure produces a change in retinal orientation only, how can the result be reconciled with the fact that, in almost all the preceding experiments, a change in retinal orientation alone has produced either slight effects or none at all? Compare, for example, those previous experiments which are comparable, in that the same 90° retinal difference is involved. Of course in those experiments, the figure which is tilted retinally is upright in the environment, whereas in this experiment its orientation to the upright is the same as that of the figure which remains normal retinally. Nevertheless, one would not have expected the retinal disorientation to produce so great a disturbance, judging by its negligible effect in these other experiments.

This difference in the effect of the two types of situations may be further clarified by the following observation. If one looks down at a drawing of a square which lies on a horizontal surface, it is clear that it becomes a diamond if it is

⁸ Rock, *op. cit.*, This JOURNAL, Experiments III to VI.

⁹ Rock, The perception of the egocentric orientation of a line, *J. exp. Psychol.*, 48, 1954, 367-374.

turned 45° . Yet, in this case it might appear that only retinal orientation is affected because there is no change with respect to environmental uprightness. If, on the other hand, one tilts his head sideways 45° from its upright position, a square which remains upright in the environment will not look like a diamond although its image is now a diamond on the retina. Granted that in this second case the unchanged orientation of the square in the environment is a major factor, it seems hard to understand why the retinal disorientation does not contribute more to the outcome in view of the role it seems to play in the first case. Why then do two situations, both of which presumably involve retinal disorientation, and that alone, lead to different results?

The following explanation suggests itself. In looking at a figure lying on a horizontal surface, we tend to organize it in terms of phenomenal directions conforming with the directions of the head. The direction from chin to forehead is 'above,' the opposite direction is 'below' and similarly for 'left' and 'right.' It has been shown elsewhere that with a given retinal orientation goes a certain egocentrically experienced orientation.¹⁰ An *O* lying on his back in a dark room perceives the orientation of a luminous line (suspended above him in a horizontal plane) in relation to the egocentric 'sides' of his head, depending upon the orientation of the line's image. It need only be added that in the situation under examination, where no part of the figure is any more 'up' or 'down' in the environment than any other part, these egocentric directions tend to take on meanings which are identical with those we speak of in characterizing environmental directions. That is to say, the egocentric 'above' is now also an 'up' direction, the egocentric 'below' is now a 'down' direction, and so forth. Thus, for example, we speak of a form inverted in this situation as 'upside down' just as we do when a form is inverted in the environment.

The results thus far suggest that retinal change, viewed as a change of the anatomical orientation of the image (and, therefore, also of the pattern of excitation in the visual cortex) is, in most instances, a factor of minor consequence. What is more important is the phenomenal orientation of a form, that is, what part is seen as up, and what part down. This orientation may be supplied by the lines of the environmental scene, by gravity, by instructions, or by recognition of the class of object shown. If, however, one or all of these are lacking, orientation may be supplied by reference to the head of the observer. In this last case, what may appear to be a matter of retinal orientation may actually be a matter of the imposition of a phenomenal orientation on the form based upon head position, a factor that unfortunately is not easily separated from retinal orientation as such.

In the case where a form remains upright in the environment, and *O* views it from a tilted position, any disturbance to recognition *would* be due to the changed retinal orientation, assuming that its phenomenal orientation is correctly given by its position in the environment. It has been shown that this factor does play some role although not a very great one in most situations. But in the case of looking at figures in a horizontal plane, the disturbance to recognition brought about by disorientation from the retinally normal position may be explained as an effect of the changed phenomenal orientation of the figure resulting from its changed

¹⁰ Rock, *ibid.*, 367-374.

orientation to the head. This would account for the difference in results in the two types of situations.

The following experiment was undertaken in an effort to give additional support to this hypothesis. According to our reasoning there should be a considerable difference, as far as recognition is concerned, produced by a 90° retinal disorientation of a Street figure seen in a horizontal plane and a 90° retinal disorientation of an upright Street figure produced by tilting the head to one side, although both situations involve the same 90° retinal shift. The latter situation yielded a score of 51% (Experiment V) in comparison with a score of 66% under optimum control conditions. For the figure in the horizontal plane we predict a much greater interference with recognition because the figure will not only be 90° retinally tilted in the anatomical sense but will also be turned 90° in relation to the egocentric up-down direction. With the figure tilted retinally *and* phenomenally, the situation ought to be comparable to the one where *O* is upright and the figure is tilted 90° in the environment.

Experiment VIII. As will be recalled, three groups of 10 *O*s each were used in Experiment IV. Two of these groups were shown, with the *O*s upright, the figures of Set A or B tilted 90° in respect to the environment. For purposes of the present experiment, they were also shown the figures of the other set on a horizontal surface, under test-conditions identical with those of Experiment VII, *i.e.*, the figures rested on the floor and *O* viewed them from above. The distance of the figures from *O*'s eyes was the same as in the preceding experiments with the com-

TABLE VI
NUMBER OF RECOGNITIONS OF COMPLEX FIGURES IN EXPERIMENT VIII
(Figures in horizontal plane)

Condition	Set A				Set B				All	%
	car	type- writer	ship	piano	bull	ele- phant	boy on bicycle	horse & car- riage		
Figures tilted 90° (<i>N</i> =10)	2	0	1	2	5	0	0	2	12	15

plex figures. All the figures were tilted 90° to the left from the retinal vertical. In every other respect the procedure was like that of Experiments IV and V.

Results. From the results shown in Table VI, it is clear that the prediction is fully verified. Only 15% of the time (12 out of a possible 80) were the figures recognized. Not only is this result considerably lower than the result of 51% in the situation where the figures are upright and *O*'s head is tilted 90° (Chi-squared gives a *P* less than $\frac{1}{10}\%$) but it is quite comparable with the situation where *O* is upright and the figures are tilted 90° (12%, see Table III). This implies that the 90° tilt in this experiment has the same detrimental effect on recognition as does a retinal *and* environmental tilt of 90° . Hence the hypothesis that disorientation in a horizontal plane involves both retinal change (in the anatomical sense) and change of phenomenal orientation, is confirmed.

SUMMARY AND CONCLUSIONS

We will first summarize the findings reported in this and the previous paper. In the first series of experiments it was shown that when *O* is presented with ambiguous figures which can be seen as either of two familiar objects depending upon which portion is seen as top, they tend to see that object which is upright in the environment in preference to the one whose image is normally oriented on the retina. The separation was brought about by requiring *O* to view the figures with head tilted 90° . The retinally normal alternative was seen so infrequently that it would seem that retinal orientation plays a minor role in this situation, although a control experiment did show that when neither alternative is favored as far as uprightness is concerned, but when one is retinally normal in orientation and the other retinally tilted, the retinally normal will be seen more frequently.

In the second series it was shown that *O* will select *as most similar*, a version of a nonsense-figure previously seen which appears upright, in preference to one which remains retinally normal. In one experiment the separation was brought about by tilting the scene 45° . The phenomenally upright figure was selected more frequently. To determine whether it was preferred because it appeared phenomenally upright or because it preserved its orientation relative to the lines of a surrounding frame of reference, a control experiment was performed in which only the test-card containing the figures was tilted 45° . The retinally normal (and truly upright) version was now preferred, thereby clearly indicating that the tilted scene in the previous experiment exerted an influence because it affected the perceived upright of space. In another experiment, in which *O* viewed the test-card with head tilted 90° , it was again shown that the environmentally upright version was chosen more frequently than the retinally normal one. A control experiment indicated that while retinal orientation played a slight role here, it was almost negligible in relation to the role of environmental orientation.

When, however, a 180° separation was created by requiring *O* to view the test-cards with inverted head, retinal orientation did play a somewhat greater role; the two versions now being selected with about equal frequency. To determine whether the upright version would be preferred with only the direction of gravity serving to define uprightness, another experiment was performed in which all lines which might serve as frames of reference were eliminated. The results again favored the upright position. This, together with the demonstration cited above that a frame affects

recognition only when it influences perception of the upright, tends to prove that the version is preferred that looks upright regardless of what cues are operating to achieve this.

In a third series of experiments, with which the present paper begins, a different technique was employed to verify the results of the preceding experiments. The test-figures were shown one at a time, *O*'s task being to identify those which looked familiar. Again it turned out that when *O* was tilted 90° , the upright versions were recognized more frequently than their tilted (but retinally normal) counterparts. In fact, a control experiment showed that when *O* was himself upright there was no appreciable superiority in recognition of the upright figures and there was no appreciable inferiority in recognition of the tilted figures. Thus this technique also shows that with figures such as these, for a 90° tilt, orientation in the environment very much influences a form's appearance, whereas orientation on the retina does not. The converse procedure in which *O* is tilted in training and upright in the test gave similar results.

In a final series of experiments fragmented figures, such as those devised by Street, were used to determine whether retinal orientation would turn out to be more important in the case of complex forms. With *O*'s head tilted 90° it was again found that upright figures were recognized more frequently than retinally normal ones. A control experiment revealed, however, that, when *O* was himself upright, recognition of upright figures was somewhat more successful and recognition of tilted figures somewhat less successful, thereby showing a significant, although quantitatively not very great, effect of retinal orientation.

When, however, *O* was told which portion of a figure was its top, a repetition of the above experiment with head tilted now indicated that the tilted figures were recognized more frequently than the upright ones. *O* is able to perceive the tilted figures in terms of their now known directions so that only the retinal factor remains to interfere with recognition of the upright ones. Additional experiments were performed to show that in the situation where *O* views figures which lie in a horizontal plane, where one might assume all directions to be equal as far as orientation with respect to environmental uprightness is concerned, *O* imposes a phenomenal orientation on figures based upon head position. Disorienting a figure in this plane will, therefore, result in a change of phenomenal as well as retinal orientation, with consequences consistent with those reported above.

There appear, then, to be two factors at work as far as orientation in form-perception is concerned. One is the perception of forms in terms

of the phenomenal directions: top, bottom and sides. It seems clear that this factor so influences phenomenal shape that a change makes for difficulty in recognition. Thus, for example, Fig. 3B of the earlier study,¹¹ has something of the shape of a mushroom in the sense of a stalk and larger top portion and this does not change with change of retinal orientation alone, as caused by tilting the head. The larger portion continues to be seen as top, the stalk as underneath and vertically oriented, etc. These aspects contribute significantly to the overall phenomenal appearance of the form. Tilt this figure in the environment 90° and it is quite another shape (unless one should happen to perceive it in such a way as to see the side as the top).

An even clearer example is the case of the square which becomes a diamond when tilted 45° , as Mach long ago pointed out.¹² As noted above, tilting the head 45° will not make an environmentally upright square look like a diamond although its image becomes a diamond on the retina. Nor will tilting the head make a diamond look like a square although its image becomes a square on the retina. We have verified this with several *O*s. Therefore, it is the selection of a phenomenal orientation, given in this case by orientation in the environment, that makes a square a square and a diamond a diamond. A square is a parallelogram having four equal sides and four right angles whose sides are parallel to the vertical and horizontal phenomenal directions, and it becomes a diamond when its diagonals assume these orientations. Evidence is lacking that retinal orientation in the purely anatomical sense has anything to do with Mach's observation.

With no information to the contrary, directions in the environment determine the phenomenal orientation of a form. To this end, since either of the two determinants of the upright alone (gravity or visual frame of reference) suffices to give these directions, either alone, or both together of course, can influence its appearance. Other factors, however, can determine phenomenal orientation, hence it need not always be given by or conform to the environment. Thus, for example, we were able to show that by informing *O* where the top of a form was, he was able to perceive it accordingly. It seems reasonable to suppose that this ability to shift the phenomenal orientation of a figure explains why we have no trouble recognizing familiar forms, such as text or a face, when tilted with the head and body. This is why, as we believe, certain observations using familiar forms such as these have led to misleading conclusions. In the case of reading the tilted book in bed, we do not see the top of the letters in the same direction as 'up' in the environment. If we did, we venture to predict

¹¹ Rock, *op. cit.*, This JOURNAL, 520.

¹² Ernst Mach, *The Analysis of Sensation*, 1914, 105-121.

that they would look quite different. Our nonsense-figures did, and they were structurally more or less equivalent to letters. Since the tilted book is viewed with head tilted, there is no retinal disorientation; and more important, familiarity with the text gives the correct selection of phenomenal orientation. There is no disturbance at all. If, however, we were to keep the book environmentally upright we would have retinal disorientation working against us and, as noted above, in the case of complex material with 90° disorientation, there would be some difficulty in recognition.

When viewing forms in a horizontal plane, environmental directions cannot serve to establish phenomenal orientation. We impose on any figure a phenomenal orientation almost as if we carry with us the directions top, bottom, and sides from the situation where we view the world from an upright posture. Top and bottom now correspond with the directions of the head. Apparently this imposition of a phenomenal orientation based upon head-position occurs only in this particular situation. If it were to happen when we view an ordinary scene with tilted head, as in many of our experiments, the vertical axis of the head would determine an 'up-down' direction in the scene. In such a situation, however, up and down are already given by the upright of the environment.

The fact that a shift in the phenomenal orientation of a form can so materially alter its appearance means that the geometrical pattern of excitation in the cortex is only a necessary aspect of the neural correlate of shape. To this pattern must be added a phenomenal top, bottom, and sides (actually the neural correlate of these subjective directions) before the shape can be said to be sufficiently determined. Whereas in the objective world we only distinguish shapes on the basis of geometrical considerations, in the phenomenal world there is also the attribute of direction. Objectively there is only a parallelogram with four equal sides and four right angles; phenomenally there is a square *or* a diamond. Of course the purely geometrical aspect of a form plays a large role. The diamond still has a great deal in common, phenomenally, with the square: these two forms obviously look more alike than a diamond and a circle or a square and a narrow, elongated parallelogram. Similarly, a change in a form's phenomenal orientation did not in our experiments always lead to a failure in recognition.

If the phenomenal orientation of a shape is abstracted, there remains, as a second factor, the orientation of the image on the retina. There is clear evidence for this factor apart from phenomenal orientation. It is true that we have not found it to be as great as we thought, but this is probably due to the fact that the disturbing effects of retinal disorienta-

tion begin to be effective only for extreme degrees of tilt and for complex or difficult figures (as shown by the increased effect with Street's figures in comparison with the other materials used). No doubt a combination of these two circumstances is more disturbing, as is suggested by the difficulty in recognizing retinally inverted words or faces.

The question arises, however, why retinal orientation, in the purely anatomical sense implied here, does influence recognition. The phenomenal equivalence of forms when they are transposed in position or size has led psychologists to conclude that the visual cortex is a more or less homogeneous medium within which the pattern of excitation and not its locus determines experience. How then is limited transposition of orientation explained? Köhler suggests that "it seems necessary to assume that the tissue of the visual center is permanently pervaded by a gradient which has a fixed direction, and which contributes to the nature of particular processes just as concretely as does the distribution of retinal stimuli in each case."¹³ Köhler was led to this view when he concluded that the effects of disorientation are due *exclusively* to changes in orientation on the retina and not to changes in an object's orientation in space.

Now that we have shown that this conclusion is incorrect, there is no need of giving up the assumption of a homogeneous medium in the visual center. The disturbance in recognition will occur without any change of retinal orientation. It would only be in the cases where the disturbance is *definitely* due to the changed retinal orientation that one might deem it necessary to subscribe to Köhler's theory. The question arises, then, whether, in these instances, the appearance of form is not phenomenally different. If it is, then it is necessary to explain the changed phenomenal shape. Since the pattern of excitation is objectively the same, only some anisotropy of the cortex, such as Köhler suggests, would seem to remain as a possible explanation.

Retinal orientation as such does not, however, seem to influence phenomenal shape. The demonstration with the square upright in the environment illustrates this point. Similarly, one would hardly say the nonsense-figures we used become phenomenally different when they remain upright and the head is tilted. Even in the case of complete inversion of the head, where retinal disorientation does definitely play a role, it does not seem correct to say that these figures are phenomenally different as far as shape is concerned. The same is probably true even of such complex material as text or script. In all these cases, if called upon to do so, one

¹³ Köhler, *op. cit.*, 29.

would probably continue to describe these forms as having the same shape when the head is tilted as when it is upright. What appears to happen when retinal disorientation alone is disturbing is that recognition is impeded or altogether prevented while phenomenal shape is more or less unchanged. The figure does appear phenomenally unfamiliar, but not different in shape. Of course, it remains possible, in certain cases, that subtle perceptual changes do occur which result in changes of phenomenal shape and thereby impede recognition.¹⁴

It would seem then that it is not change of phenomenal shape as a result of retinal disorientation which causes the difficulty but rather it is difficulty in access to the trace which results in strangeness in appearance or lessened familiarity. The question of whether a disoriented figure is hard to recognize is, functionally speaking, a question of the relationship of a trace and an ongoing perceptual process. When the shape is phenomenally changed, this leads to difficulty in access to the trace, since in general traces are contacted by similarity of the process underlying the percept to trace. This is no doubt the nature of the disturbance in the case of changed phenomenal orientation. In the case of changed retinal orientation, however, the difficulty in access to the trace may be due to a difference between the anatomical orientation of the process underlying the percept and of the trace. It is plausible to suppose that traces preserve their original orientations. If true, this fact would be important in connection with the general problem of trace-localization. Whether traces are localized in the region where the original perceptual process takes place has been a question of interest. Since recognition does not require identity of locus, some psychologists falsely imply that the trace is not localized. Logically, the trace may be localized and recognition take place regardless of where the new process occurs. Recently, evidence has been advanced to show that communication with a trace nearer to the new percept is more apt to occur than communication with a trace of a more distant percept.¹⁵ This suggests that the trace *is* localized at the place of the original perceptual process, or at least that that place is in some way involved. The present results similarly suggest that traces may preserve their original *orientations* or at least that the process whereby a new percept arouses a similar trace must be one in which the anatomical orientation of the original perceptual process plays a role.

¹⁴ Other perceptual effects having to do with direction, as for example the vertical-horizontal illusion, may reveal that subtle anisotropies of the visual center do exist which might influence shape in certain ways, but it remains to be established that these effects are themselves due to the retinal factor.

¹⁵ Hans Wallach and Pauline Austin, Recognition and the localization of visual traces, this JOURNAL, 67, 1954, 338-340.

As noted above, the orientation of the trace does not matter much unless the form is complex or difficult and the difference in orientation considerable. To a great extent the orientation can be changed without impairing recognition just as can the position or size of the figure. In cases of successful recognition of retinally disoriented figures, trace-arousal clearly depends upon phenomenal similarity. Whereas recognition frequently fails when an incorrect phenomenal orientation prevails (even if retinal orientation remains unchanged), it generally succeeds when the phenomenal orientation is such as to lead to a similar phenomenal shape (even if retinal orientation is changed). Thus the phenomenal attribute takes precedence over identity of retinal-cortical orientation as a factor in trace-arousal in these cases and they may well be the rule rather than the exception. In the instances where retinal orientation as such does influence recognition, there is a definite limitation of equivalence as far as recognition is concerned. This limitation, we have argued, is not due primarily to anisotropy of the visual center.

In other words, it is suggested that in those cases where a form is not recognized or looks strange and unfamiliar as a result only of retinal disorientation, it is not due to change in phenomenal shape but to inadequate communication with the appropriate trace. This, in turn, is due to difference in orientation of trace and process which for some as yet unexplained reason impairs access to the trace. Generally, this only becomes of consequence in the case of complex or difficult forms and orientation changes of 90° or more. We have been able, however, to show that retinal disorientation *can* play a role even with relatively simple figures when the influence of phenomenal orientation is held constant. Thus, other things being equal, a sensitive test may show that arousal of a trace having the same orientation as a present perceptual process is more likely to occur than arousal of one having a different orientation.

Since in all the experiments reported here recognition is involved, we may assume that there is a competition between the two factors for trace-arousal. The test-figure correctly perceived in terms of phenomenal orientation is more likely to communicate with the trace because it remains phenomenally unchanged *qua* shape; the retinally normal test-figure is more likely to do so because it has the same orientation, anatomically, as the trace. The outcome in any particular instance depends upon factors which influence these two determinants such as the characteristics of a figure in relation to angle of tilt, in the case of phenomenal shape; and degree of tilt and degree of difficulty or complexity in the case of anatomical orientation.

RESPONSE-SEQUENCES AND THE HYPOTHESIS OF THE NEURAL QUANTUM

By ULRIC NEISSER, Harvard University

The quantal hypothesis put forward by Békésy, and amplified by Stevens, Morgan, and Volkmann, asserts that sensory magnitude along a given dimension, say, auditory loudness, is not indefinitely divisible.¹ Any increase in the loudness of a tone must be of at least a certain small amount, the so-called 'neural quantum,' and represents the activation of an additional neuron or group of neurons. The size of the quantum, as measured by the maximal amount of additional external stimulation needed to produce an increase in loudness, is assumed to remain roughly constant in the listener during an average experimental session. Because of the constantly varying sensitivity of every *S*, however, identical stimulation does not excite the same number of quanta from moment to moment, and behavioral variability can be expected.

It has been shown that, if experimental conditions are appropriate, and if *S* adopts an efficient criterion (*i.e.* reports only changes of two or more quanta), these assumptions lead to two predictions concerning the psychometric function for differential loudness-judgments.² The function should be rectilinear (more precisely, should rise linearly from zero to one in a certain range, remaining at zero below this range and at one above it), and it should have a slope such that the weakest stimulus-difference that always produces a positive response has precisely twice the value of the strongest difference that always produces a negative response. Of the two predictions, the first is at variance with both the cumulative ogive of the classical phi-gamma hypothesis and other modern psychophysical theories;³

* Received for publication March 30, 1956. This paper is based on a dissertation submitted to Harvard University in partial fulfillment of the requirements for the Ph.D. degree. The study, made at the Psycho-Acoustic Laboratory, Harvard University, was supported by contract with the Office of Naval Research (U.S. Navy, Project NR142-201, Report PNR-183). The author is now at Brandeis University.

¹ Georg van Békésy, Ueber das Fechner'sche Gesetz und seine Bedeutung für die Theorie der akustischen Beobachtungsfehler und die Theorie des Hörens, *Ann. Phys.*, 7, 1930, 329-359; S. S. Stevens, C. T. Morgan, and John Volkmann, Theory of the neural quantum in the discrimination of loudness and pitch, this JOURNAL, 54, 1941, 315-335.

² Stevens, Morgan, and Volkmann, *op. cit.*, 318.

³ E. G. Boring, Urban's tables and the method of constant stimuli, this JOURNAL, 28, 1917, 280-293; W. P. Tanner and J. A. Swets, A decision-making theory of visual detection, *Psychol. Rev.*, 61, 1954, 401-409.

the second is unique, in that no other theory predicts less than an infinite range of psychometric variability. An imaginary set of data satisfying the quantal predictions is shown in Fig. 1.

The quantal hypothesis does not require that every psychometric function display these characteristics, but only those functions measured under conditions that combine maximal ease of observation with close temporal contiguity of standard and comparison stimuli. The procedure typically used for this purpose is that of Stevens, Morgan, and Volkman, herein-after called the 'quantal method.' The details of this method will be given later. Here we need only mention its most controversial characteristic: stimuli are presented in series of 25, and all the stimuli comprising any particular series are identical. With this technique, the predictions of the quantal hypothesis have been verified several times.⁴

A number of experimenters have explicitly declined to accept the quantal hypothesis, chiefly on either of two grounds. In the first place, several attempts to replicate the early findings have failed entirely or partially.⁵ While such negative findings need not be considered conclusive, a theory that can be confirmed only half the time cannot be regarded as satisfactory. In the second place, Blackwell and others have argued, without regard to this negative evidence, that the rectilinear functions may be spuriously generated by the quantal method.⁶ The argument hinges on the sequential dependence among *S*'s responses, and before discussing it in detail we will consider briefly such dependencies in general.

There have recently been a number of demonstrations that successive responses in psychophysical experiments are not independent; runs of responses of 'yes' and 'no' occur more extensively than a binomial model would permit. Collier has pointed out that there are two fundamentally different mechanisms that might create perseverative tendencies of this kind.⁷ On the one hand, *S*'s sensitivity may fluctuate, producing a run of positive responses when it stays high and a run of negative responses when it stays low. (This assumes that the same stimulus is being presented repeatedly.) The fluctuations may be spontaneous and independent of experimental conditions (the only case that Collier considers), or they may depend on such experimental variables as warning lights or intertrial intervals. Changes in sensitivity are a fundamental postulate of the quantal hypothesis, as of most other psychophysical theories. Consequently, unless the changes are rapid enough to leave successive responses uncorrelated, perseverative effects can be expected.

⁴ Stevens, Morgan, and Volkman, *op. cit.*, 315-335; G. A. Miller and W. R. Garner, Effect of random presentation on the psychometric function: implications for a quantal theory of discrimination, this JOURNAL, 57, 1944, 451-467.

⁵ B. M. Flynn, Pitch discrimination; the form of the psychometric function and simple reaction-time to liminal differences, *Arch. Psychol.*, 280, 1943, 1-41; J. F. Corso, The neural quantum in discrimination of pitch and loudness, this JOURNAL, 64, 1951, 350-368.

⁶ H. R. Blackwell, The influence of data-collection procedures upon psychophysical measurement of two sensory functions, *J. exp. Psychol.*, 44, 1952, 306-315.

⁷ George Collier, Intertrial association at the visual threshold as a function of intertrial interval, *J. exp. Psychol.*, 48, 1954, 330-334.

On the other hand, it is also possible that the outcome of each trial depends causally on the outcome of the preceding trial (or trials). Such a mechanism, which we may call a *response-linkage*, would again make successive responses interdependent. Any response-linkage would distort the psychometric function from what we might like to consider its 'true' form, since the response-probabilities would depend not only on the relation of the stimulus to momentary sensitivity but also on a basically extraneous variable—the preceding response. The simplest possible response-linkage is that described mathematically by the Markov process, in which the dependence extends only to the immediately preceding response, and to no earlier ones. Even in such a model, substantial distortions of the function can occur. The author has shown elsewhere how a plausible type of Markov process would distort a basically quantal psychometric function into one that appeared to have an ogival shape.⁸

We may now discuss Blackwell's suggestion. He argues that the quantal method, in which *S* is aware that all the stimuli in each series are alike, favors a complex type of response-linkage which he calls "channelization." It is said to be characteristic of this process that a predominance of 'yes' responses at the beginning of any series will sharply raise the probability of 'yes' responses in the remainder of the series, as *S* strives for consistency. This would tend to sharpen the upper corner of an ogival function, since high probabilities of 'yes' would be made even higher. A similar mechanism, *mutatis mutandis*, would sharpen the lower corner, and the result would be a spuriously rectilinear shape. The slope of the function is not predicted.

The status of this counter-hypothesis has been complicated recently by Jerger's report that successive responses in the quantal method are, in fact, independent.⁹ As we have seen, this finding is difficult to reconcile with the quantal hypothesis itself, or with any theory that postulates fluctuations of sensitivity.

This paper is a report of experiments relevant to the following questions: (1) What difficulties are encountered in attempting to confirm the predictions of the quantal hypothesis? (2) Can response-interdependence be demonstrated in the quantal method? (3) If so, is the interdependence due to fluctuations of sensitivity, spontaneous or otherwise, or to a response-linkage—and, in the latter case, is the linkage like a Markov Process or is it more complex? Are several of these mechanisms working simultaneously?

In the first preliminary experiments, the straightforward quantal method was used repeatedly. In some later sessions, various kinds of catch-trials were introduced into the supposedly homogeneous series to clarify the nature of the response-interdependence. Finally, in the main experiment, two new independent variables were introduced. First, a warning light was

⁸Ulric Neisser, A methodological study of the quantal hypothesis in auditory psychophysics. Ph.D. Dissertation, Harvard University, 1955, Ch. 1.

⁹J. G. Jerger, On the independence of successive responses in the quantal psychophysical method, this JOURNAL, 68, 1955, 145-147.

sometimes presented preceding the stimuli, to discover whether, as is commonly supposed, such signals achieve some control over sensitivity. Secondly, the interval between successive stimuli was varied to provide information about the time-dependence of the sequential effects.

METHOD

Subjects. One woman and three men were employed as Ss. All were students between the ages of 18 and 21 yr., and naïve with respect to the purpose of the experiments. For reference purposes, a rough measurement of each S's absolute threshold for tones of 1000~ was carried out before the experiment began, using a modified method of limits. The (monaural) thresholds were all within the normal range. Each S received extensive practice with the quantal method (at least eight full sessions) before the beginning of the experiments reported here. In the practice periods, the importance of accuracy was stressed; S was told to avoid false reports as completely as possible, while responding to every stimulus heard. This instruction was dramatized by the occasional presentation of a 'series' without stimuli. If S responded on such an occasion, he was rebuked. After a few practice sessions, positive responses never occurred in these series. It was hoped that this procedure would lead to the establishment of a two-quantum criterion of judgment.

Procedures. The experiments were conducted in two adjacent sound-shielded rooms provided with an intercom system. E and the apparatus were in one room, S in the other. S's room was dark. He sat in a comfortable easy chair, wearing PDR-10 earphones in Grason-Stadler semiplastic CZW-6 cushions, designed for optimal shielding of extraneous noise. S's comfort was a major consideration, and he was instructed to report any source of distraction, however minute. If he did so report, the ongoing series (or, if necessary, the entire experimental session) was discarded. Furthermore, S was encouraged to rest whenever he wished.

A 1000~ tone was presented steadily to S's right ear. The sensation-level of this tone ranged from 50 to 80 db., but was never changed during any one session. The experimental stimuli consisted of increments to the magnitude of the background-tone. The increments were 300 m.sec. long, with rise- and decay-times of 7 m.sec. Twenty-five successive increments formed a series.

In the straightforward quantal method, all the increments in any given series were identical. They were spaced 3 sec. apart. S was to depress his response-key each time he heard one. No light or other warning signal preceded the individual stimulus-increments. An experimental session typically included four series (100 judgments) at each of 8 to 12 stimulus-magnitudes, and lasted about 2 hr. including rests. Two Ss participated in the routine quantal experiments, JG contributing six sessions and GR seven.

The same two Ss (JG and GR) served in the experiments involving 'catch-trials.' In these sessions, unknown to S, the 25 stimuli comprising a series were not necessarily all the same. In some of the sessions, the stimulus-increment was simply omitted at a few fixed positions in certain series. This procedure, called *null series*, was used several times with each S. Another procedure, called *interchange series*, was used with GR only. In each such session, E selected one strong and one weak value of the stimulus-increment, both within the range of response uncertainty. He

presented series made up of 20 stimuli of one type and 5 of the other at fixed positions. Thus either weak stimuli were concealed in series made up primarily of strong stimuli, or vice versa.

In the main experiment, the interval between stimulus-increments was varied. First, preliminary work was done at various intervals with the same two Ss as before, using no warning light (seven sessions). Then two new Ss (*ED* and *LP*) served in a balanced design as follows: four interstimulus-intervals were employed (1.5, 2.5, 4.0, and 6.0 sec.). Each S had 16 sessions, 2 per day, and was exposed to each interval 4 times—twice with and twice without a warning light. Order-effects were counterbalanced. The light preceded each stimulus by 0.75 sec. In other respects the routine quantal method was followed, except that series of only 15 stimuli were used with the 6-sec. interval to keep the total duration of the series within bounds. During each session, 120 to 125 judgments were recorded at each stimulus-value.

Apparatus. The stimulus-increments were produced and controlled by a variable-shunt circuit essentially like that of Miller and Garner¹⁰. A graphic recorder provided a sequential record of stimuli and responses. A detailed description of the apparatus is available elsewhere.¹¹

RESULTS

Form of the function. For each experimental session, an empirical psychometric function was formed by plotting $p(y)$, the relative frequency of 'yes' responses (key-presses), against stimulus-magnitude. The latter was measured in terms of the Weber ratio, $\Delta p/p$, where p is the constant sound-pressure level of the background-tone and Δp the increment in sound-pressure.

The majority of these functions do not have the rectilinear shape predicted by the quantal hypothesis, nor any other specific shape. A few roughly rectilinear functions and a few ogives occur, but most of them are jumbles of points with no clear character except for a roughly monotonic increase.

Thresholds. For purposes of analysis, each empirical psychometric function was fitted visually with a straight line. All points for which $p(y)$ lay between 0.03 and 0.97 were equally weighted in the fitting process, while points outside these limits were not used at all. Independent repetition of the fitting process showed it to be highly reliable. (Independent *DLs* correlated at 0.99, intercept ratios—see below—at 0.88.)

The place where the line of best-fit crossed 0.50 was taken as the *DL* (see Fig. 1). While the values of these thresholds are not specifically

¹⁰ Miller and Garner, *op. cit.*, 454.

¹¹ Neiser, *op. cit.*, Ch. 2.

relevant to the purposes of this study, two points are worth noting:

- (1) The thresholds are of the same order of magnitude as those obtained by other experimenters with the quantal method. At a sensation-level of 50 db., the liminal sound-pressure ratio is about 0.055—roughly 0.5 db.
- (2) The thresholds are systematically smaller when a warning light is used than when it is not used. Analysis of variance shows that this effect is reliable at the 1% level of confidence. This strongly suggests that *S*'s sensitivity (or 'attention') is not constant, and that its fluctuations can be partially controlled by the warning signal.

Slopes. The upper and lower intercepts, where the best-fit line crosses the ordinates 1.00 and 0.00 respectively, were noted. The quantal hypothesis predicts that the upper intercept will have twice the value of the lower. Consequently, the ratio of these two stimulus-values was formed in every case (see Fig. 1). The mean *intercept ratio* in the 52 psychometric functions available is 2.15, quite close to the predicted 2.00. The distribution of intercept ratios is positively skewed, the first quartile falling at 1.85, the median at 2.08, and the third quartile at 2.40.

The balanced design of the main experiment permitted a four-dimensional analysis of variance (*S* X warning light X interval X replication). Such an analysis was carried out on the logarithmic transforms of the intercept ratios. No variable and no interaction reaches significance at the 5% level. Furthermore, the correlation between *DLs* and intercept ratios is 0.14, a value not reliably different from zero.¹² This constancy of the intercept ratio may also be regarded as a confirmation of the quantal hypothesis.

First-order sequential effects. Two measures were used in the evaluation of sequential effects: (1) conditional probabilities, and (2) autocorrelations. The notion of conditional probability may be clarified by an example. Consider the fictitious sequence of responses YNNNNYYYYYNNNNNNYNYNYYYYYYY. (Here Y refers to a 'yes' or key-press, N to a 'no' or failure to press the key). In this sequence $p(y)$ is 14/25. We obtain an estimate $p_y(y)$ of the conditional probability that 'yes' will occur after an immediately preceding 'yes' by examining the 13 responses which follow a Y. In the example, $p_y(y)$ has the value 9/13. Similarly, we form an estimate $p_o(y)$ of the conditional probability that a 'yes' will occur just after a 'no'; $p_o(y)$ is 4/11 in this instance.

¹² This does not mean that the threshold is uncorrelated with the *geometric slope* of the function, which represents behavioral variability (*i.e.* the size of the quantum). Indeed, these two quantities must covary if the intercept ratio is to remain fixed; in the present experiment they were correlated at 0.94.

Values of $p_y(y)$ and $p_o(y)$ were calculated for every psychometric point in the main experiment. (By psychometric point is meant the total of all the series measured at a given stimulus-value on one S in a single session.) These values were plotted as functions of stimulus-magnitude. The resulting 'conditional psychometric functions' are comparable to the ordinary psychometric function which plots $p(y)$ against the stimulus. For convenience, the conditional psychometric functions were visually fitted with straight lines. A pair of these functions is illustrated in Fig. 1.

A perseverative tendency in the responses is indicated whenever $p_y(y)$ is higher than $p_o(y)$. In fact, by a well-known theorem, the first-order

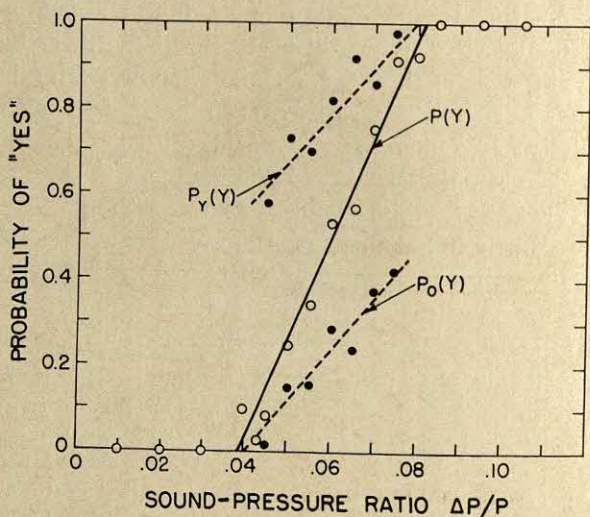


FIG. 1. A FICTITIOUS PSYCHOMETRIC FUNCTION ILLUSTRATING VARIOUS QUANTITIES. The open circles represent values of $p(y)$; the solid line is the best visual fit; the solid dots are the corresponding values of $p_y(y)$ and $p_o(y)$; the dotted lines are the conditional psychometric functions drawn through them. The DL is 0.061; the intercept ratio is 2.13; the autocorrelation at threshold is 0.55.

autocorrelation is given exactly by the difference between the two.¹³ In the example above, the first-order autocorrelation is $(9/13) - (4/11)$, or 0.33.

The application of these measures to the data of the present experiments reveals powerful perseverative effects. The conditional psychometric function for $p_y(y)$ invariably lies well above that for $p_o(y)$. The con-

¹³ G. A. Miller, Finite Markov processes in psychology, *Psychometrika*, 17, 1952, 156, 154.

sistency of the perseverative effect is suggested by the following 'sign test': there are 176 psychometric points in the main experiment at which both $p_v(y)$ and $p_o(y)$ lie between 0.03 and 0.97. In these cases, $p_v(y)$ is the larger of the two 175 times, and smaller once!

It is convenient to be able to represent the degree of sequential interdependence in an entire experimental session with a single number. The first-order autocorrelation *at the threshold-stimulus* is a useful statistic for this purpose. The threshold-stimulus may be found from the ordinary

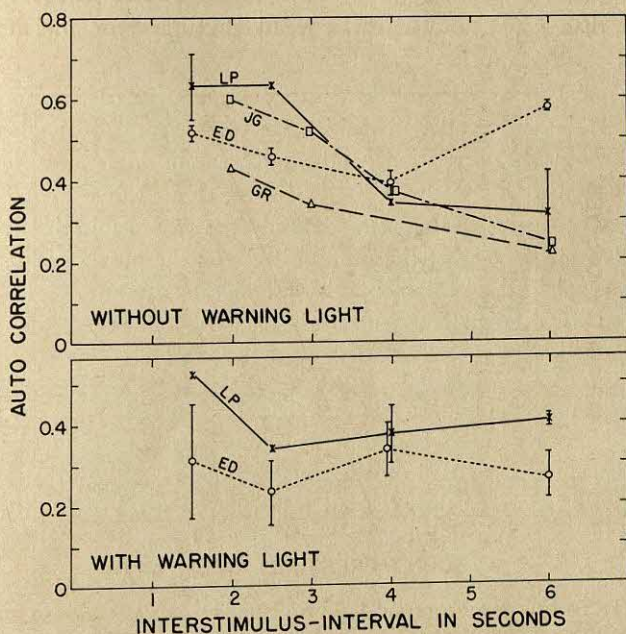


FIG. 2. FIRST-ORDER AUTOCORRELATIONS AT THE THRESHOLD AS FUNCTIONS OF INTERSTIMULUS-INTERVAL

Vertical lines indicate the ranges of the paired replications.

psychometric function for $p(y)$. The values of $p_v(y)$ and $p_o(y)$ at this stimulus can then be estimated from the conditional psychometric functions. The difference of these two conditional values is the desired autocorrelation at the threshold. The autocorrelations at the threshold in the main experiment are exhibited as functions of interval between and stimuli of the warning light in Fig. 2. It is evident that the autocorrelation is systematically high and positive, ranging from 0.15 to 0.71.

The autocorrelations at the threshold were subjected to Fisher's Z-

transformation, and an analysis of variance like the one already described was performed on the transformed scores. The analysis indicates that the presence of the warning light significantly reduces the autocorrelation (at the 1% level). No other variable or interaction reaches the 5% level of confidence.

Fig. 2 includes the results of the preliminary experiments at various rates. When these are combined with the data of the balanced design, there is some suggestion of a tendency for the autocorrelation to decline with increasing interstimulus-interval. This effect is not revealed by the analysis of variance because of the paradoxical rise for *ED* at the longest interval.

Although these analyses demonstrate strong interdependence among responses, they are not decisive with respect to the cause of the dependence. A decline in autocorrelation as the interstimulus-interval is lengthened might be expected equally on the basis of a slow change in sensitivity or a linkage of successive responses; indeed, it is remarkable that the decline is so slight. The effect of the warning light on the perseveration can likewise be explained plausibly by either hypothesis. Accordingly, we must go on to more detailed treatments of the data.

Higher-order sequential effects. Conditional psychometric functions for dependencies extending over more than one step may be calculated by a method analagous to that described above for first-order dependence. These functions are based on the conditional probabilities that a response of 'yes' will occur on the n th trial following a 'yes,' or following a 'no.' Higher-order autocorrelations at the threshold may then be obtained as before. These have been computed for lags (values of n) of two, three, and four in every session of the main experiment.

As might be expected on any interpretation, the higher-order autocorrelations are typically smaller than the first-orders. The decline is slow. Detailed presentation of these data appears unnecessary; only the medians of the autocorrelations of the first four orders are presented in Table I.

Autocorrelations to 10 steps were computed directly from the response-sequences at a few psychometric points. Occasionally these remain strikingly high; correlations of 0.30 at lag 10 are not uncommon.

These results permit certain inferences about the underlying process. The conception of an uncomplicated one-step linkage of responses must be ruled out. By an established theorem about a Markov Process, the n th order autocorrelation in such a linkage is precisely the n th power of the

first-order autocorrelation.¹⁴ A glance at Table I shows that this model is inapplicable here. We are therefore dealing either with a linkage of responses extending over many steps, or, more plausibly, with a slowly changing sensitivity.

An altogether different type of analysis supports this conclusion. Each psychometric point is made up of a number of individual series, measured at various times during the experimental session. If the level of sensitivity

TABLE I
MEDIAN OF AUTOCORRELATIONS AMONG RESPONSES TO A STIMULUS AT THRESHOLD,
LAGS ONE THROUGH FOUR, MAIN EXPERIMENT
($N=39$)

Lag	Median r
1	0.41
2	0.35
3	0.32
4	0.25

changes slowly during the session, these several series may sample a number of different levels. This possibility was checked at every psychometric point in the main experiment at which the overall $p(y)$ lay between 0.40 and 0.60. The frequency of the responses of 'yes' in each component series was compared with an expected frequency based on $p(y)$ for the psychometric point as a whole. A chi-squared test with appropriate $df.$ was performed against the null hypothesis that the several frequencies were samples from one population.

Forty-two chi-squared tests of this kind were made. In 27 of them, the null hypothesis could be rejected at the 1% level of confidence. This striking lack of homogeneity among series measured at single stimulus-values points clearly to the occurrence of large drifts of sensitivity.

Length of runs. A run of responses was defined as a sequence of one type of responses uninterrupted by the other type. The series illustrated above has five runs of 'yes,' of varying lengths, and four runs of 'no.' The lengths of runs may be expected to be related to the period of the underlying fluctuations. Therefore, frequency-distributions of lengths were compiled, separately for runs of 'yes' and 'no,' and for each interstimulus-interval. The tabulation was made for both Ss in the balanced design of the main experiment; they do not differ appreciably and have been combined in Fig. 3.

¹⁴ Miller, *op. cit.*, 154.

Several aspects of Fig. 3 are worth noting. (1) Some very long runs occur. These are related to the high long-term autocorrelations already described. (2) If we consider runs of length one, *i.e.* isolated responses, we find that at each interstimulus-interval 'no' appears more frequently than 'yes.' Altogether, there are 238 isolated responses of 'no's' and only 168 isolated responses of 'yes's.' This suggests a form of rapid fluctuation in which momentary lapses of sensitivity (or attention) are more common than momentary peaks. (3) The curves for the four interstimulus-intervals

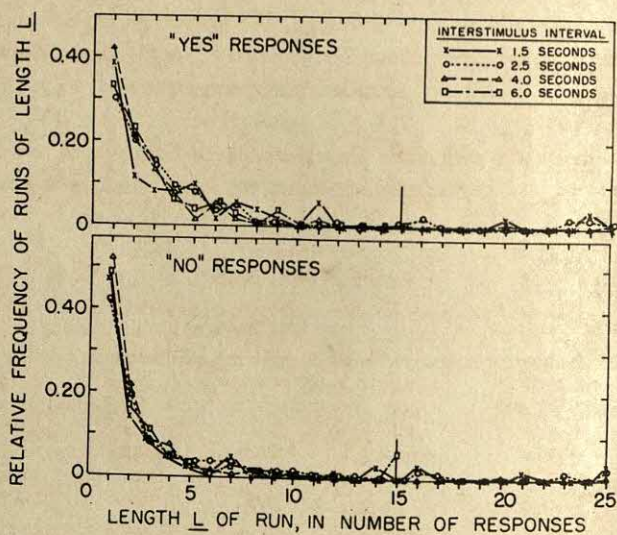


FIG. 3. DISTRIBUTIONS OF RUN-LENGTHS FOR RUNS OF RESPONSES OF 'YES' AND OF 'NO'

Data of *ED* and *LP* are combined; the several curves are for different interstimulus-intervals.

lie very close together. This coincidence is remarkable, since a run of length N at the 6-sec. interval lasts four times as long as does a run of length N at the 1.5-sec. interval. It appears that the passage of time itself is irrelevant to the fluctuations; instead, these are controlled by the interstimulus-interval.

Catch-trials. The results reported so far have all pointed to changes of sensitivity as the cause of the response-perseveration. The experiments with catch-trials, however, provide evidence for the existence of linkage of responses as well.

We will first consider those series in which the stimulus was occasionally omitted. The *Ss* were rarely fooled by these *nulls*. In fact, $p(y)$ on the null trials is only 9/289, or 0.03. Thus, there remain 280 null trials on which the response 'no' was made. It is instructive to compare the conditional probability $p_{o.n}(y)$ of a response of 'yes' on the trial *following* one of the 280 nulls with the ordinary conditional probability $p_o(y)$ that a 'yes' will occur on the trial *following* a normal 'no.' In the case of a one-step linkage of responses, these two probabilities should be identical, since the preceding response is 'no' in each case. In a pure fluctuation of sensitivity, however, they should be very different. The responses of 'no' to the null indicates nothing whatever about *S's* sensitivity, so the trial following the null is simply a randomly sampled trial; consequently, $p_{o.n}(y)$ should be equal to $p(y)$.

The results of the null trials are presented in Table II. It appears that the response of 'no' to the null has about the same effect on the next trial

TABLE II
RESULTS OF NULL SERIES

The $p_{o.n}(y)$ Is the Probability of a Response of 'Yes' to a Real Stimulus on the Trial After a Response of 'No' Was Made to a Null Stimulus
(The number of judgments on which each probability-estimate is based is given in parentheses.)

S	$p(y)$ real stimuli	$p(y)$ null stimuli	$p_y(y)$ real stimuli	$p_o(y)$ real stimuli	$p_{o.n}(y)$
JG	.61 (1336)	.01 (139)	.77 (718)	.42 (420)	.37 (137)
GR	.39 (1600)	.05 (150)	.69 (547)	.20 (833)	.16 (142)
Totals	.49 (2936)	.03 (289)	.73 (1265)	.27 (1253)	.27 (279)

as does any other 'no.' We are therefore dealing with some degree of response-linkage.

Similar considerations apply to the *interchange series*. In a one-step linkage of the Markovian type, the conditional probabilities are fixed; $p_y(y)$ and $p_o(y)$ should be unaffected by the concealed or open nature of the stimulus-presentation. But since a weak stimulus will more frequently follow a 'yes' when it is presented concealed in a series of strong stimuli than when it is presented openly, the Markov model predicts changes in $p(y)$. The $p(y)$ should be higher for concealed than for open weak stimuli, and lower for concealed than for open strong stimuli.

Precisely the opposite predictions follow from the notion of a drifting

sensitivity in the interchange series. The $p(y)$ should not be affected by the manner of presentation, since all presentations sample the sensitivity randomly. As a consequence, the conditional probabilities in concealed and in open presentation should differ substantially.

The results of the interchange series are presented in Table III. The manner of presentation has no great effect on the conditional probabilities,

TABLE III
RESULTS OF GR INTERCHANGE-SERIES, ALL SESSIONS COMBINED
(The number of judgments on which each probability-estimate is based is given in parentheses.)

Type of stimulus	$p(y)$	$p_o(y)$	$p_c(y)$
weak, presented openly	.28 (500)	.66 (153)	.11 (304)
weak, concealed in strong series	.47 (120)	.64 (81)	.10 (39)
strong, presented openly	.68 (480)	.87 (293)	.34 (163)
strong, concealed in weak series	.44 (125)	.97 (34)	.26 (86)

but markedly alters $p(y)$. This again suggests the reality of a causal linkage between one response and the next.

Channelization of responses. It seems desirable to evaluate Blackwell's hypothesis that extended linkages of responses tend to distort ogives into straight lines. A fully adequate test is impossible in the absence of rectilinear functions, but one may profitably compare the psychometric points which steepen the functions with those that flatten them.

Points were defined as falling in the upper corner of their psychometric function when they represented a $p(y)$ between 0.89 and 0.99, and as belonging to the lower corner if $p(y)$ lay between 0.01 and 0.11. Now, points in the upper corner and lying to the *left* of the best-fit line, as well as points in lower corner lying to the *right* of the line, were classed as *rectilinear points* because they tend to steepen the function. On the other hand, upper points to the right of the best-fit line and lower points to its left were termed *ogival points*, since they would more nearly fit a classical ogive. In Fig. 1 there are three rectilinear and two ogival points. (Since Fig. 1 has been drawn to represent a good quantal function, these points deviate less than is typically the case.)

A chi-squared test was applied to the hypothesis that rectilinear points would be associated with high autocorrelations and ogival points with

lower autocorrelations in the main experiment. This test is exhibited in Table IV. The hypothesis is not substantiated; indeed, the non-reliable trend runs in the other direction. So far as we can determine from the available data, then, response-interdependence does not spuriously straighten out ogival functions.

TABLE IV

CHI-SQUARED TEST OF RELATIONSHIP BETWEEN TYPE OF DATA POINT AND
MAGNITUDE OF FIRST-ORDER AUTOCORRELATION IN THE MAIN EXPERIMENT
(Chi-squared equals 1.29 with $df = 1$.)

	Ogival points	Rectilinear points	Marginals
$r(1)$ above median	22	7	29
$r(1)$ below median	18	11	29
Marginals	40	18	$N = 58$

DISCUSSION

Interdependence of responses. The results of these experiments demonstrate a very powerful perseverative interdependence among successive responses in the quantal method. We have seen that this is to be expected in nearly every psychophysical theory, and it is in agreement with the findings of many workers using other procedures. It is in apparent conflict with the data of Jerger.¹⁵ The discrepancy may be ascribed in part to procedural differences,¹⁶ and in part to the small samples on which Jerger's serial correlations are based. He presented each S with a single series of 20 increments at each stimulus-magnitude. Long-term effects, like those found here, would not only be undetectable in such short sequences but would necessarily bias the evaluation of short-term correlations by preventing genuine random time-sampling.

Nature of the interdependence. The data show clearly that several distinct mechanisms are contributing to the perseveration in responding.

To some extent, the outcome of each trial exerts a direct influence over the outcome of the next. This linkage is demonstrated by the results with catch-trials. The response of 'no' to a null depresses $p(\gamma)$ on the next

¹⁵ Jerger, *op. cit.*, 145-147.

¹⁶ The interstimulus-interval used by Jerger (5.5 sec.) is overlapped by the intervals used here, and the unusually long rise-times and durations of his increments have been explored experimentally by the author and do not seem to be relevant. Two remaining procedural differences may, however, be important: Jerger used much lower sensation-levels than those in this study, and his S s were not trained in the same fashion.

trial about as much as does an ordinary 'no.' In the interchange series, the dependence of each response on its predecessor is not much affected by the magnitude of the stimulus that elicited the predecessor.

On the other hand, fluctuations of sensitivity also occur. Not only do the high-order autocorrelations decay much too slowly to fit the Markov model, but a chi-squared test shows that $p(y)$ to a particular stimulus-value does not remain fixed during the experimental session. Both of these results suggest slow fluctuations, but some fast changes in sensitivity are present as well. Isolated 'no' responses occur much more often than isolated 'yes' responses.

It is important to note that the changes of sensitivity do not occur independently of experimental events. "Spontaneous fluctuation" is probably a myth. Two lines of evidence require this conclusion. First, the threshold (like the autocorrelation) is lower when a signal light is employed than otherwise, indicating that the light partially controls the S's sensitivity, as indeed it should. Secondly, the data on run-length prove that the sensitivity-drift depends on the rate of stimulus-presentation. If this were not so, the distribution of lengths of runs should be strictly determined by the interstimulus-interval. In fact, the interstimulus-interval has no effect at all on the distributions. We must conclude that the fluctuations are not influenced by the passage of time in itself, but rather by the occurrence of stimuli, and perhaps of stimulus-series.

The origins of patterning of response in the quantal method, and by implication throughout psychophysics, are therefore exceedingly complex. They include both linkages of responses and stimulus-bound changes of sensitivity.

Validity of the quantal hypothesis. The evidence of this study is ambiguous with respect to the quantal hypothesis. The constancy of the intercept-ratio near the quantal value 2.00 must be regarded as a confirmation, since all values between one and infinity are equally likely on the basis of any other existing theory. The absence of rectilinear functions is less encouraging, but need not be regarded as disproof. The lack of rectilinearity may be due in part to the distorting influence of Markov processes. Furthermore, the present Ss were less sophisticated than most of those used in previous validations of the quantal hypothesis, and they may also have been less strongly motivated. Indeed, it is possible that these factors made them more prone to linkages of responses. In any event, it appears unlikely that the rectilinear functions found by other experimenters have been due to the "response-channelization" proposed by

Blackwell. With the present *Ss*, the data that steepen the psychometric function are not characterized by higher autocorrelations than those that flatten it.

SUMMARY

Experiments were conducted with the so-called quantal psychophysical method, in which *S* is required to detect as many as possible of a series of identical increments to the intensity of a steady tone. In preliminary work, the quantal method was followed routinely a number of times. In other sessions, the magnitude of the increment was sometimes changed during the series, without *S*'s knowledge. In the main experiment, the interval between successive stimulus-increments was systematically varied from session to session, as were the presence and absence of a warning light preceding each increment.

The results indicate that successive responses in the quantal method are not independent; a strong perseverative tendency is present. Various lines of analysis show that the origin of the interdependence of responses is complex. It is due in part to a direct causal linkage between successive responses, and in part to *S*'s changing sensitivity. Both fast and slow changes in sensitivity occur. It is shown that these fluctuations do not occur spontaneously in time but are themselves controlled by the inter-stimulus-interval, by the length of the series, and by the warning light.

The data are regarded as predominantly favorable to the quantal hypothesis concerning sensory discrimination, since the psychometric functions obtained have approximately the predicted slope. Their failure to display rectilinearity may be due to extraneous factors.

THE EFFECT ON VISUAL PERCEPTION OF STIMULATING THE BRAIN WITH POLARIZING CURRENTS

By GARTH J. THOMAS and PETER A. STEWART, University of Illinois College of Medicine

The field-theory of cortical integration was proposed originally by Köhler and Wallach to explain figural after-effects and other perceptual phenomena.¹ The theory supposes that cortical neurons involved in the representation of a figure on the cortex generate electrochemically potential differences between themselves and surrounding tissues. As a result, ionic currents are presumed to flow through these tissues with a distribution which depends on the particular figure as well as the particular location in the cortex. It is also postulated that during continued flow of these currents their distribution changes because of alterations in the resistance of membrane barriers through which current passes (satiation). Perceptual phenomena are supposed to parallel isomorphically these electrochemical phenomena.

Köhler and his associates were the first to study physiological implications of the theory by recording slow or steady potential changes from various electrode locations on the scalp.² Recording from the occiput, they found slow potential changes which were correlated with the movement of a visual stimulus across the visual field. More recently they have found similar slow potential changes associated with auditory stimuli.³

A second procedure for testing physiological implications of the field-theory was first reported by Lashley, Chou and Semmes.⁴ These investigators implanted a criss-cross of gold foil strips over the surface of the visual cortex of one group of monkeys, and inserted short gold pins perpendicularly into the visual cortex of another group. If the distribution of currents in the cortex is significant in perceptual organization, these metallic conductors should presumably disturb visual perception by distorting the normal patterns of current-distribution. Their presence

* Received for publication April 23, 1956. From the Department of Psychiatry, Neuropsychiatric Institute, Chicago, Illinois. This investigation was made possible by a research grant (M-694) from the National Institute of Health, U. S. Public Health Service.

¹ Wolfgang Köhler and Hans Wallach, Figural after-effects, an investigation of visual process, *Proc. Amer. phil. Soc.*, 88, 1944, 269-357.

² Wolfgang Köhler and Richard Held, The cortical correlate of pattern vision, *Science*, 110, 1949, 414-419; Wolfgang Köhler, Richard Held, and D. N. O'Connell, An investigation of cortical currents, *ibid.*, 96, 1952, 290-330.

³ Wolfgang Köhler, W. D. Neff, and Jonathan Wegener, Currents in the auditory cortex of the cat, *J. cell. comp. Physiol.*, Suppl. No. 1, 45, 1955, 1-24; Köhler and Wegener, Currents in the human auditory cortex, *ibid.*, 45, 1955, 25-54.

⁴ K. S. Lashley, K. L. Chou, and Josephine Semmes, An examination of the electrical field theory of cerebral integration, *Psychol. Rev.*, 58, 1951, 123-136.

failed to cause any detectable defects in visual recognition, figure-ground differentiation, closure of a figure from dot patterns, discrimination of color, form or distance, visual orientation, or visual reaction-time. A similar study was reported by Sperry, Miner, and Myers.⁵ They implanted a thick net of tantalum pins in the visual cortex of cats. Here also, the cats continued to learn and retain elaborate visual discriminations, some of which depended on very fine acuity. In a related study, Sperry and Miner inserted thin mica plates perpendicularly into the visual cortex of cats.⁶ Resulting deficits in visual discrimination were attributed to the amount of neural tissue damaged rather than to interference with distribution-patterns of figural currents.

Another implication of the theory is that externally applied direct currents flowing through the cortex should interfere with the normal distribution of figural currents, cause changes in satiation, and hence result in perceptual changes. Köhler and Wallach reported a preliminary test of this implication, but they did not pursue the work because of untoward effects, such as clouding of the visual field and persisting headache.⁷ Lashley, Chou, and Semmes also discussed this method of testing the Köhler theory, but rejected it as not being a critical test since almost any theory of neural organization would have to predict an effect on cortical function of direct currents of sufficient magnitude.⁸ As they pointed out, the Köhler theory does not specify the magnitude of current density in the cortex required to produce perceptual effects. Their point is well taken, but few theories in psychology or biology are so rigorously and quantitatively formulated that they can be unequivocally tested by single crucial experiments. The series of experiments reported below was therefore undertaken with a more empirical than critical intent to see whether any perceptual changes could be induced by passage of practical amounts of direct current (8-10 ma.) through the head via scalp electrodes. Positive results would have been consistent with Köhler's theory, although they could not have been interpreted as proving it in any critical sense. The negative results actually found at least show a range of variables over which predictions from the theory are not borne out.

Similar remarks apply to the relevance of the present data to the theory of Libet and Gerard, which also attributes integrative function to slowly changing currents spreading through the brain.⁹ They suppose that spreading currents would

⁵ R. W. Sperry, Nancy Miner, and R. W. Myers, Visual pattern perception following subpial slicing and tantalum implantations in the visual cortex, *J. comp. physiol. Psychol.*, 48, 1955, 50-58.

⁶ Sperry and Miner, Pattern perception following insertion of mica plates into visual cortex, *J. comp. physiol. Psychol.*, 48, 1955, 363-369.

⁷ Köhler and Wallach, *op. cit.*, 329.

⁸ Lashley, Chou and Semmes, *op. cit.*, 126-127.

⁹ Benjamin Libet and R. W. Gerard, Steady potential fields and neuron activity, *J. Neurophysiol.*, 4, 1941, 438-455.

function by modifying thresholds of cortical neurons through which they pass. Externally applied direct currents should therefore affect cortical function, but implications regarding perceptual phenomena *per se* would be only very general.

METHODS OF APPLYING THE ELECTRICAL STIMULUS

Current from a 45-v. dry cell battery was sent through a simple circuit consisting of a potentiometer and rheostat with a voltmeter and ammeter so placed in the circuit that the potential and current flowing between the electrodes could be determined. The electrodes were thin silver disks, $7/8$ in. in diameter. They were held in place on the head by large rubber bands that encircled the head.

Experience showed that passage of the current was less painful when the skin resistance was low. For this reason the areas on the scalp where the electrodes were to be placed were first washed with detergent solution, then soaked with saline solution for a few minutes and finally electrode paste was massaged into the skin before the electrodes, liberally smeared with paste, were applied.

Except for the first experiment reported below, the electrodes were always placed in the same location, one on the occipital protuberance and the other on the vertex. Both were in the midline. Preliminary investigation showed that passage of the current with the electrodes placed to any extent bilaterally from the midline resulted in a marked galvanic vestibular response consisting of vertigo, nystagmus, and sometimes feelings of nausea. In all but one of the experiments, both directions of the stimulating current were tried, *i.e.* inion positive relative to vertex and vice versa.

Only one session with current on was run on any one day because of the possibility that any satiating effects of the current might be relatively long lasting. Control runs were taken either before observations were made with stimulating current on, or on days when no current was used.

S controlled the onset of the current by turning the potentiometer, increasing the current slowly at a rate determined by his tolerance for the stinging pain in the skin under the electrodes. After the current was on for a short time the pain usually subsided and became negligible. For each run with current on, *S* turned the current up slowly to as much as he could comfortably stand, usually between 8–10 ma. Immediately after the perceptual tests, the voltage and current values were noted, and *S* then turned off the current gradually by returning the potentiometer to the zero position.

EXPERIMENTS AND RESULTS

Experiment 1: Visual equation of lengths of lines. A luminous line 40 cm. long and 2 mm. wide was bisected by an opaque panel which could be moved on slides. The panel divided the luminous line into two horizontal lines, end to end, and separated by 10 cm. ($3^{\circ} 22'$ angular subtense). A fixation-point was provided in the center of the panel in line with the horizontal lines. With the panel centered and *S* gazing at the fixation-point, two luminous horizontal lines, each subtending 5° visual angle, were seen in an otherwise dark field and located symmetrically to the

right and left of the fixation-point. *S* could move the panel back and forth by turning a hand wheel, thus reciprocally changing the lengths of the lines. His task was to adjust the two lines to apparent equality in length while maintaining fixation with the preferred eye. *S*'s head was held steady and at a constant distance from the stimuli by means of a chin and forehead rest.

Measurements were taken under five conditions: a control with no current passing through the head, and four experimental conditions during which the stimulating current was on. The experimental conditions consisted of two electrode placements, 1 in. to the right or left of the midline on theinion, the other electrode being on the vertex in both instances. For each location measurements were taken under two conditions of polarity, vertex positive or negative relative to the occipital electrode. At a given session electrodes were put in place and 10 control settings of apparent equality of length of the lines were obtained. After a short rest period, *S* turned up the stimulating current and a series of five settings of equality were made. Only one of the four experimental conditions was used on a given day. Control measures were always obtained first. *E* offset *S*'s previous setting alternately to the right or left.

It was thought that the right or left location of the occipital electrode would result in a different current density in the two occipital poles which would supposedly result in different levels of satiation. If such a state of affairs obtained it might be inferred from Köhler's theory that changes in the metric of visual space would result. Objects would appear smaller in the visual field on the side represented in the more highly satiated hemisphere.

Data from two *Ss* (*GT* and *RS*) are presented. *GT*, the senior author, was of course aware of the purpose and issues involved in the experiment; *RS*, however, was naïve with regard to these matters.

Table I presents the results. The direction of the average errors (difference between the point of subjective equality (*PSE*) and true equality in millimeters) is indicated by algebraic signs before the numbers. Positive errors indicate increased apparent size of *S*'s left visual field relative to the right, *i.e.* the stimulus in the left visual field was the smaller. The reverse holds for negative errors. For one *S* one of the differences between control and experimental average errors (no current vs. current on, occipital electrode on the left, negative) is statistically significant at the 1% level of confidence (*t*-test). All other relevant differences fail to reach significance. It is obvious that there is no relation that is significant and consistent for both *Ss* between direction or magnitude of the average errors and the location of the electrodes or the polarity of the current.

Experiment II: Critical flicker frequency. The field-theory of cortical action makes no specific interpretation of the phenomenon of critical flicker frequency (*CFF*). It might be argued, however, that an increased

overall level of cortical satiation presumably induced by the polarizing currents might have the effect of decreasing cortical sensitivity, an effect which might plausibly be thought to lower the *CFF*.

The stimulus was a circular spot of light 0.5 in. in diameter, seen in an otherwise dark field. At a viewing distance of 1 m. it subtended about 45' visual angle, thus falling well within the fovea. A tiny black spot of ink in the center of the milk glass aperture served as a fixation-point. A series wound motor rotated a sectored

TABLE I

AVERAGE ERRORS AND SD IN MILLIMETERS OF JUDGED EQUALITY OF LENGTH OF LINES AS A FUNCTION OF ELECTRODE-LOCATION AND POLARITY OF STIMULATING CURRENTS

(Positive errors indicate over-estimation of the line in the left visual field; negative errors, over-estimation in the right visual field.)

S and location and polarity of electrodes	Av. error	SD	Electrical stimulation			
			before tests		after tests	
			ma.	volts	ma.	volts
RS						
Control, no current	+1.3	2.6	—	—	—	—
Electrode to right, vertex pos.	-0.8	4.7	8.0	6.0	8.3	6.0
Electrode to right, vertex neg.	-2.1	5.0	8.0	5.5	8.0	5.5
Control, no current	-2.2	5.6	—	—	—	—
Electrode to left, vertex pos.	-9.1	4.0	8.1	6.7	8.8	6.0
Electrode to left, vertex neg.	-5.8	2.7	8.0	6.2	8.3	5.1
GT						
Control, no current	+2.8	2.4	—	—	—	—
Electrode to right, vertex pos.	-0.2	5.6	8.0	6.0	8.3	6.0
Electrode to right, vertex neg.	+3.8	4.0	8.0	7.0	8.6	5.0
Control, no current	+2.1	4.4	—	—	—	—
Electrode to left, vertex pos.	-0.7	2.9	8.0	10.0	9.4	9.0
Electrode to left, vertex neg.	+1.3	2.3	8.2	7.2	8.5	7.0

disk ($L/D = 1$) which interrupted a beam of light at a point of focus. The light, which transilluminated the milk glass aperture, came from a lamp powered by a voltage regulating transformer. The operating level of the lamp was calibrated each time it was turned on by adjusting its intensity to a given value of the output of a photoelectric cell. A series of neutral density filters provided steps of flash luminance about one log unit apart. The speed of the motor rotating the sectored disk was controlled with a rheostat and measured with a magneto tachometer. The optical system was calibrated with a Macbeth illuminometer.

The *CFF* of two *Ss* (*GT* and *RS*) was measured after 15 to 20 min. of dark-adaptation at five levels of flash luminance, five determinations at each level always progressing from lower to higher levels of flash-luminance. Only critical flicker thresholds were taken, *i.e.* flash rate was set well above *CFF* and then lowered slowly until *S* reported the detection of flicker. After completion of these control measures (no current), *S* was again allowed 15-20 min. rest and dark-adaptation and then the current was turned on and kept on until *S* completed another set of

25 threshold-determinations, five at each of the five levels of flash-luminance. The current was always set at 8–10 ma. at the beginning of the experiment. Only one electrode placement and one direction of current-flow was used in this experiment. The positive electrode was on the vertex and negative electrode on theinion, both in the midline. The procedure was replicated three times on three different days, and as there are no systematic differences among the data from different days, the results for all three experiments were combined, making a total of 15 judgments per point.

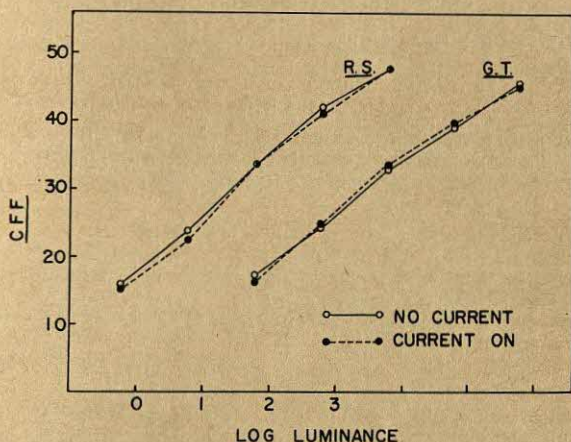


FIG. 1. CFF CONTOURS OBTAINED FROM RS AND GT WITH AND WITHOUT POLARIZING CURRENTS

The curves for GT have been displaced two log units to the right.

TABLE II

MEAN CFF ($N=15$ JUDGMENTS) WITH SD AS A FUNCTION OF LOG LUMINANCE (MILLILAMBERTS) WITH AND WITHOUT APPLICATION OF POLARIZING CURRENTS

Log flash luminance	RS				GT			
	Control (no current)		Experimental (current on)		Control (no current)		Experimental (current on)	
	CFF	SD	CFF	SD	CFF	SD	CFF	SD
-0.235	16.0	0.63	15.1	0.68	17.3	1.61	16.3	1.23
+0.749	23.8	1.22	22.3	0.78	24.2	1.35	24.5	0.96
+1.779	33.6	2.47	33.4	1.58	33.0	1.27	33.6	1.29
+2.787	41.9	1.62	41.1	1.10	39.1	1.33	39.5	2.36
+3.776	47.7	0.70	47.8	0.57	45.5	1.49	45.2	2.59

Results are shown in Fig. 1 and Table II. Table III shows the characteristics of the stimulating currents. The differences between the average CFF obtained during electrical stimulation and the control conditions are,

with one exception for each *S*, within one flash per second of each other. No consistent effect of the current on the average *CFF* or on the variability of the thresholds is apparent. It is clear that 8–10 ma. flowing through the head from scalp electrodes has no effect on the photopic flicker contour.

Experiment III: Absolute threshold for form. It is a reasonable inference from Köhler's theory that a high level of satiation induced by polarizing currents through the cortex would not only cause distortion of figural processes but also, because of increased resistivity of the medium,

TABLE III

CURRENT IN MILLIAMPERES AND VOLTAGE ACROSS THE ELECTRODES AT THE BEGINNING AND END OF EACH OF THE THREE EXPERIMENTAL SESSIONS DURING WHICH *CFF* WAS BEING MEASURED

Session	RS				GT			
	before tests		after tests		before tests		after tests	
	ma.	volts	ma.	volts	ma.	volts	ma.	volts
1	8.0	9.0	7.9	8.5	8.0	5.0	8.1	5.0
2	8.0	6.0	7.8	6.3	8.0	2.5	7.6	4.0
3	8.0	7.5	7.9	8.0	8.1	8.0	8.4	8.5

larger potentials would be required to set up a given figural process. If so, one would predict that passage of current through the brain would raise the absolute threshold for form-recognition. To test this possibility, absolute thresholds for recognition of the orientation of a T-pattern were measured in terms of luminance of a tachistoscopically presented flash, with and without passage of polarizing currents through the head.

The target consisted of a T-pattern made of black Scotch tape stuck on a circular milk glass window. The window subtended $7^{\circ} 46'$ visual angle at a viewing distance of 55 in. The stem and cross bar of the T were of equal length, $4^{\circ} 10'$ subtense and $47'$ visual angle in width. The milk glass disk could be rotated so the T-pattern could be set in each of four orientations, cross bar at the top, bottom, right or left. The window was transilluminated by brief flashes presented by a simple rotary disk tachistoscope similar to that described by Schlosberg and Woodworth.¹⁰ Duration of the flashes is not known but it is reasonable to assume they were relatively constant. Light from an automobile head lamp operated by step down and voltage regulating transformers was brought to a point of focus in the plane of the rotary shutter. The voltage across the bulb was monitored by means of a voltmeter. Beyond the point of focus the light beam diverged through fixed filters, wedge and balancer and fell on the milk glass window, illuminating it uniformly. A small red cross

¹⁰ R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 1954, 92

situated $12^{\circ} 48'$ visual angle above the center of the T-pattern served as a fixation-stimulus. *S* viewed the target with his preferred eye. His head was held steady in a chin and forehead rest. Thresholds were always taken after at least 35 min. of dark-adaptation.

To measure thresholds as quickly as possible, a simplified method of limits, descending series only, was used. *E* set the flash luminance to a level at which *S* could clearly discriminate the orientation of the T-pattern and then he presented a series of flashes in which the flash-luminance was decreased in steps of 0.2 log units until *S* made his first mistake in calling the orientation of the T-pattern. The luminance of the flash on which he made his first error constituted the threshold. After a little practice the method yielded consistent thresholds suitable for the purpose at hand but undoubtedly somewhat higher than would have been obtained with more elaborate psychophysical methods.

Ten determinations of threshold by two *Ss* were made under three conditions: no current (control series always came first), current on, vertex positive and inion negative, and vice versa. Only a single series with current on was done on any one day.

Table IV shows the means and standard deviations of the thresholds in millilamberts obtained under the various control and experimental conditions as well as the characteristics of the stimulating current. Clearly there

TABLE IV
ABSOLUTE THRESHOLDS FOR FORM-PERCEPTION (MEANS AND SD OF 10 MEASURES IN MILLILAMBERTS) WITH AND WITHOUT APPLICATION OF POLARIZING CURRENTS

S and experimental condition	Thresholds		Electrical stimulus			
	mean	SD	before tests		after tests	
			ma.	volts	ma.	volts
RS			—	—	—	—
Control (no current)	1.21×10^{-5}	6.75×10^{-6}	8.0	7.6	8.4	7.0
Current, vertex pos.	1.07×10^{-5}	5.35×10^{-6}	—	—	—	—
Control (no current)	1.11×10^{-5}	4.70×10^{-6}	8.0	9.0	8.8	8.4
Current, vertex neg.	1.23×10^{-5}	5.71×10^{-6}	—	—	—	—
GT			—	—	—	—
Control (no current)	1.01×10^{-5}	4.24×10^{-6}	8.0	6.0	7.8	6.4
Current, vertex pos.	1.14×10^{-5}	3.75×10^{-6}	—	—	—	—
Control (no current)	1.13×10^{-5}	3.84×10^{-6}	8.2	10.3	9.7	8.5
Current, vertex neg.	0.94×10^{-5}	2.36×10^{-6}	—	—	—	—

are no significant shifts in the threshold for form-perception attributable to the presence or polarity of currents up to 8 ma. applied to the brain through scalp electrodes.

Experiment IV: Reversible perspective. Köhler has interpreted the phenomenon of reversible perspective in terms of his field-theory of corti-



cal function.¹¹ In terms of isomorphism, each of the alternate perceptual organizations of ambiguous figures is considered to be represented by different figural processes in the cortex. Persistence of one of the perceptual alternatives sets up a flow of currents around that cortical pattern resulting in satiation which, in turn, tends to block or weaken the figural process until finally the weakened organization is supplanted by the other perceptual organization. The deduction from this theory that continued observation of an ambiguous figure should result in an increasing rate of apparent reversal has been verified by Brown's results.¹² If passage of polarizing currents through the brain has the effect of raising the general level of satiation, it should also cause an increase in reversal-rate. Some preliminary observations with the Necker cube yielded suggestive but equivocal results, so a more elaborate experiment was performed using a moving stimulus-pattern similar to that described by Brown.¹³

A disk 5 in. in diameter was rotated at 15 r.p.m. in the horizontal plane by a constant speed electric clock motor. Around the rim of the disk were six equally spaced vertical rods, $\frac{1}{8}$ in. in diameter and 4 in. high. This apparatus was placed just behind a 5.5×3.5 in. window in an opaque screen. The window was covered with white paper and was transilluminated by an automobile head lamp situated 5.5 ft. behind the screen. The arrangement produced very clear and sharp shadows of the revolving pins on the paper in the stimulus-aperture. An ink spot in the center of the window served as a fixation-stimulus. Usually one sees a pattern of vertical lines moving alternately to the right or left in a circular motion in three dimensions as in Lissajous figures. Sometimes, however, the three dimensional pattern breaks down and one sees the vertical lines moving back and forth in the two dimensional frontal plane. The Ss were instructed to report each time there was a change in direction of apparent rotation in three dimensions and each time there was a change from three dimensional movement to two dimensional and vice versa. The various modes of appearance often described of the patterns of movement in the frontal plane (a perceptually unstable situation) were lumped together in one category of two dimensional movement patterns. S reported the occurrence of observed changes in perceptual organization of the pattern by pressing a key which activated an electric counter.

At any one session the number of perceptual changes was recorded during three 3-min. periods of observation separated by 2-min. intervals of rest during which time the stimulus-light was off. Ambient lighting in the room was low but objects were clearly visible. Number of changes was recorded under three conditions: a control condition with no current and two conditions with current on, vertex positive and vertex negative relative to the electrode on the inion. Observations under one condition only were made on any one day. The sequence of three sessions was repeated

¹¹ Köhler, *Dynamics in Psychology*, 1940, 1940, 67-82.

¹² K. T. Brown, Rate of apparent change in a dynamic ambiguous figure as a function of observation-time, this JOURNAL, 68, 1955, 358-371.

¹³ Brown, *op. cit.*, 360-363.

in reverse order, so the reported rates of reversal are based on the mean of two 3-min. periods of observation. Current was applied in the experimental sessions during the first 3-min. period only. Three Ss were used, *GT* and two (*AK* and *LO*) who were naïve with regard to the purpose of the experiment.

The results are shown in Table V. Table VI shows the characteristics of the stimulating currents. The previous findings of Brown's that continued

TABLE V
REPORTED CHANGES PER MINUTE IN PERCEPTUAL ORGANIZATION OF AN AMBIGUOUS
STIMULUS WITH AND WITHOUT APPLICATION OF POLARIZING CURRENTS
(Periods of electrical stimulation are indicated by asterisks.)

S	Session (3 min. each)	No current (control)	Current	
			vertex pos.	vertex neg.
LO	1	12.3	9.5*	9.0*
	2	13.5	13.8	11.2
	3	11.9	13.2	13.7
AK	1	7.8	8.3*	6.7*
	2	9.4	9.8	10.5
	3	10.3	10.8	10.5
GT	1	8.2	8.5*	8.7*
	2	9.0	9.0	8.5
	3	9.5	8.8	10.0

TABLE VI
VOLTAGE AND CURRENT APPLIED DURING PERIOD OF ELECTRICAL STIMULATION WHILE
Ss REPORTED CHANGES IN PERCEPTUAL ORGANIZATION OF
AN AMBIGUOUS STIMULUS

S	Session	Polarity (vertex pos.)				Polarity (vertex neg.)			
		before tests		after tests		before tests		after tests	
		ma.	volts	ma.	volts	ma.	volts	ma.	volts
LO	1	8.0	18.0	7.0	18.0	10.0	12.0	9.6	12.8
	2	10.0	13.0	13.5	11.0	10.0	13.8	8.5	15.0
AK	1	9.0	14.0	8.0	15.0	10.0	6.0	9.8	6.5
	2	10.0	5.0	10.0	5.0	10.0	10.0	13.0	9.0
GT	1	9.4	18.0	9.0	18.0	10.0	13.0	10.0	13.0
	2	10.0	11.0	11.0	10.3	9.8	18.0	7.5	19.0

observation of ambiguous figures tends to result in increased rate of change are verified.¹⁴ With few exceptions the frequency of perceptual change increases with successive periods of observation. No effect, consistent among all three Ss, is, however, related to the passage of current in

¹⁴ Brown, *op. cit.*, 358-371.

either direction between the electrodes during the actual period of electrical stimulation (indicated by entries with asterisks in Table V) or in the subsequent periods of observation of that session.

DISCUSSION

The results reported above clearly indicate that application of the maximal practical steady current to the conscious human brain has no detectable effects on visual perception. The meaning of this fact depends primarily on the magnitude of the current actually flowing in the visual cortex under the experimental conditions. Indications that physiologically significant amounts of current flowed through the head were obtained incidentally from two observations. When the scalp electrodes were placed more than a few inches bilaterally from the midline, strong vestibular responses were obtained during current flow, as already noted. It was also observed that if currents as low as 3 or 4 ma. were turned on or off suddenly, flashes of light were clearly seen. This response may well have depended on retinal stimulation rather than stimulation of intracranial tissues, but even so, it indicates effective magnitudes of current-flow at locations remote from the electrodes.

In an attempt to achieve a more quantitative estimate of actual current magnitudes inside the head, appropriate measurements were made on and in the skull of a cat anesthetized with nembutal. Polarizing currents were applied through disk electrodes on the shaved skin of the head. Both electrodes were on the midline, one just above the eyes, the other on the back of the head, over the lamboidal ridge. Potential measurements were made with a pair of chlorided silver wires with just the tips exposed. One wire, used as the reference electrode, was kept on the skin adjacent to the posterior polarizing electrode. The other electrode was placed as desired by means of the electrode holder of the stereotaxic instrument in which the cat's head was held. The measurements obtained indicated that roughly one-third of the total current actually flowed through intracranial tissue, and that this flow was spatially rather homogeneous; no steep potential gradients were detected inside the skull. Extrapolating from the volume of the cat brain to that of the human, it was possible from these measurements to estimate that the current-density in the brain under our experimental conditions, with 8-10 ma. applied to the scalp electrode, was on the order of 10 μ a. per mm.², corresponding to a potential gradient of about 10 mv. per mm.

These admittedly rough estimates are of the same order of magnitude as the threshold (rheobase) values for square-wave stimulation of excised peripheral nerve trunks.¹⁵ Their meaning in terms of stimulation of the cortical neurons depends of course upon the micro-geometry of current flow in the cortex, as well as on the

¹⁵ Routine measurements on excised frog whole sciatic nerve trunks in a student laboratory yield rheobase values of about 50 mv. and 20 μ a. between Ag-AgCl electrodes 5 mm. apart.

accommodation properties of cortical neurons, since the experimental currents were turned on gradually rather than abruptly. They do suggest, however, that the currents flowing through the cortex were of sufficient magnitude that the negative results obtained can not be simply dismissed as due to completely inadequate stimulation. On one occasion, *GT* submitted for several minutes to a polarizing current of 16 ma., an intensity twice that used in most of the experiments. The stronger current resulted in a severe stinging pain in the skin under the electrodes, and in a very intense deep headache which subsided as soon as the current was turned off. During the period of intense electrical stimulation, no changes could be observed in the visual field as *S*'s gaze wandered among the familiar objects on the laboratory table in front of him.

The failure of such currents to produce any change in visual perception suggests several interesting possibilities. In the first place, it may well be that the effective resistance of glia and tissue fluid in the brain is much lower than that of neurons, so that, to put the physical situation in very simple terms, the current flows mainly around rather than through excitable structures. Such a possibility would of course be completely incompatible with the assumptions of both the Köhler and the Gerard-Libet theories of functionally important current-flow in the cortex. Another equally plausible possibility is that cortical neurons are actually traversed by the current, but that they accommodate so very rapidly and completely that their normal function continues undisturbed. This possibility also is hardly consistent with the postulates of either of the above mentioned theories.

A third possibility, which is not in itself inconsistent with either the theories of Köhler or of Libet and Gerard, is that what is important in the cortex is not stimulation of individual cortical neurons *per se*, but the presence of relatively steep potential gradients across layers of many neurons. This possibility can best be tested by stimulation in animals with chronically implanted electrodes, arranged in a variety of ways so as to produce a range of potential gradients in the cortex. It is difficult to see, at the present state of our knowledge, how this latter possibility can be true unless the first possibility above is also true, since steep potential gradients imply high current-density in a conducting system, and we have already seen that the current densities produced in this experiment are probably of the order of magnitude of threshold for nerve fibers.

Apart from these possible implications of our results, their direct bearing on the Köhler theory is not clear-cut. At first glance, they would seem to argue strongly against this theory, since if it is true that our currents are near threshold, they must surely have comparable effects to the currents he postulates as being functionally associated with a figure process in the

cortex. There are, however, a great many details involved in the electrochemistry and volume circuitry of his postulated figure processes which have not yet been (and probably cannot for some time be) quantitatively outlined or even approximated, hence it is not quite fair to try to predict in any quantitative sense what the interactions between applied currents and perceptual phenomena will be on the basis of his theory. It is surely a little surprising, however, that no interaction at all is found.

SUMMARY

A physiological implication of the field-theory of cortical function was investigated by passing 8-10 ma. of direct electric current through the head, via scalp electrodes, while various perceptual functions were tested. The electrodes were located on the vertex and occipital protubance and both directions of current flow were usually tested. The perceptual tests were visual equation of length of lines, *CFF*, absolute threshold for form, and frequency of perceptual change of an ambiguous figure. No effects of the current on any of the perceptual tests could be detected. Estimates based on extrapolations from appropriate measures of current flow in a cat's head indicate that current density in the brain of the *Ss* was probably of the order of magnitude of threshold values for nerve fiber. It is unlikely that the negative results can be attributed to inadequate electrical stimulation. Consequences of the results for Köhler's field-theory or for the theory of Libet and Gerard are equivocal because their present state of formulation does not allow quantitative predictions of the effects of applied currents on perceptual phenomena. The purpose of the experiments was more empirical than critical.

SEMANTIC GENERALIZATION AT THE LEVEL OF THE CONDITIONING EXPERIMENT

By ALBERT A. BRANCA, Veterans Administration Hospital,
Fayetteville, N.C.

Semantic generalization and mediated generalization have recently been proposed as useful concepts in the study of language.¹ The present experiment attempts to use these concepts in a study of language in persons without speech. We sought to test the generalization of a conditioned response (CR) to materials that were semantically related to the conditioned stimulus (CS). Control experiments with normal subjects (Ss) failed to reveal such generalization. This result led to a reexamination of the studies on semantic generalization and to the conclusion that there is at present no acceptable evidence of the existence of semantic generalization in a *conditioning experiment*. Accordingly, this experiment was designed to test whether, in the delimited area of laboratory conditioning, a response attached to a word or picture would generalize to a semantically related but otherwise dissimilar stimulus-object. Four situations were provided in which the occurrence of semantic generalization could be detected: (1) A response attached to the picture of an object was tested by the presentation of the name of that object; (2) a response attached to the name of an object was tested by presentation of the picture of the object; (3) a response attached to one word was tested by the use of a semantically related word that was different in spelling and appearance; and (4) a response that had been attached to three different but semantically related words was tested with a fourth having essentially the same meaning but different spelling and appearance. The stimuli used are shown in Table I.

METHOD

Subjects. Eighty-eight students (76 men and 12 women) participated in this study. Their mean age was 21 yr.

Apparatus. The stimulus-objects were images from 35-mm. slides. The letters

* Received for publication September 9, 1955. This paper is a modification of a portion of a doctoral dissertation submitted to the University of North Carolina. This investigation was directed by Dr. Irvin S. Wolf. The author is now at the University of Delaware.

¹ C. N. Cofer and J. P. Foley, Jr., Mediated generalization and the interpretation of verbal behavior: I Prolegomena, *Psychol. Rev.*, 49, 1942, 513-540. Gregory Razran, Experimental semantics, *Trans. N. Y. Acad. Sci.*, Series II, 14, 1952, 172-177.

of projected words were in capitals, 5 in. high. The frame of the pictures was 7 in. $\times$ 16 in. square. Electric shock was used as the unconditioned stimuli (*US*). Responses were recorded by a galvanometer with a continuous ink-writing recorder.

Procedure. Each *S* appeared at the experimental room at a prearranged time. He was encouraged to sit as comfortably as possible in a deep leather chair and to rest for 5 min. while *E* checked the apparatus and prepared to record responses to questions which served to establish *S*'s naïvete about this experiment.

After *S*'s remarks were recorded, the stimulating electrodes were attached to his right ankle, and the galvanometer electrodes were fitted against the dorsal and

TABLE I
EXPERIMENTAL STIMULI USED IN TESTS

	Conditioned stimuli	Generalized test-stimuli
Test I	picture of a hat	HAT
Test II	HAT	picture of a hat
Test III(a)	FREEZE	CHILL
Test III(b)	BROOK	STREAM
Test IV	CREEK RIVER STREAM	BROOK

palmar surfaces of *S*'s right hand. A ground wire was attached to the right ankle above the stimulating electrodes.

After all the adjustments had been made, *S* was again encouraged to relax. He was informed of the harmless but painful nature of the shocks he would receive during the experiment and then permitted to feel a mild shock. After this initial experience with shock, which reassured most *Ss*, he was invited to "see how much shock you can stand. I will stop turning it up when you feel you cannot stand anymore." The shock was then turned on and its strength slowly advanced until *S* indicated at what point *E* should stop. The stimulator was then turned off and *S* was again encouraged to relax.

After the shock-level had been determined, the galvanometer and its recorder were turned on and adjusted. The following instructions were then read to *S*.

Instructions: In just a moment I shall project some slides on the screen in front of you, but before we start I wish you to get as comfortable as possible. If you feel like moving or stretching, do so now for while I show you the slides I want you to remain as quiet and relaxed as possible. You will not be required to say or do anything except watch the screen intently. Try to maintain an attitude of relaxed attention, as though you were reading a good book, comfortable but alert and attentive to what is being shown to you on the screen. When the slides appear concentrate on them but do not say anything aloud or even move your lips.

The slides I am about to show you are mostly words in black and white like this. [A word was projected.] There may also be some pictures in color like this. [A picture was projected.] Occasionally, and under certain conditions, the electric shock which you have already experienced, will be turned on for a brief period, like this. [The shock was turned on briefly. The sensitivity of the galvanometer was set in accordance with the extent of the resultant, unconditioned *GSR*.]

Conditioning of the *GSR* to the word or picture was then begun. A list of pre-

sumably neutral words was shown *S*. Each slide was projected for about 5 sec. which was estimated by counting. The period between slides was allowed to vary from 10 to 25 sec. Each slide, except the conditioned slide, was shown to *S* only once. The conditioned slide was projected at irregular intervals during the training and at each presentation it was accompanied by shock throughout the last second that it remained on the screen. Training was terminated when the conditioned stimulus appeared to evoke a greater *GSR* than control stimuli in the training list but before the controls had lost their power to evoke *GSRs*. Training was also prolonged until *S*'s skin-resistance assumed a relatively stable level.

At the conclusion of training and without change in general procedure the test-slides, composed of seven neutral slides and one slide that was semantically related to the conditioned slide, were presented to *S*.

At the end of the test-series the conditioned stimulus, without reinforcement, was projected on the screen. Putting this test for conditioning at the end of the test for generalization instead of at the end of the conditioning runs constituted a minor but important departure from the usual procedure. A positive conditioning-test in this situation indicated with some certainty that *S* had been conditioned at the time of the test for generalization. In earlier experiments it had been possible to argue that failure of generalization was due to the unreinforced test of conditioning which extinguished the conditioned response.

After about 5 sec. the conditioned slide was removed and after 20 sec. the recorder and the galvanometer were turned off and the room lights were turned on. *S* was told that the experiment was over but that he was to remain seated for a few additional moments to answer some questions about the experiment.

It was the purpose of this post-experimental interview to determine to what extent *S* had discovered the purpose of the experiment and if he had formed any hypotheses about the relationships between the conditioned stimulus, the stimulus-testing generalization, and the unconditioned stimulus. It was also the object of this interview to determine whether *S* had seen and been aware of the presence of the test-stimulus in the test-list. *S* was also asked if he expected to receive shock and under what specific circumstances.

Usually the desired information followed the first and most general of a group of increasingly specific questions. If it did not, other questions were asked until the information was elicited. A brief recall- and recognition-test was administered to assure *E* that *S* was aware of the occurrence of the experimental stimuli. At the conclusion of the post-experimental interview, *S* was released from the apparatus. He was asked not to discuss the experiment with others and then he was dismissed.

The procedures in the four experimental situations were identical except for the choice of stimuli. In the fourth situation, however, the simultaneous conditioning of three stimuli necessitated certain changes. During the training period, the three words, RIVER, STREAM, and CREEK, were distributed by chance throughout the training list. Presentation of any of the three conditioned words was always followed by shock during the training. The test for conditioning which followed the showing of the test-list involved the presentation, without reinforcement, of one of the conditioned words. The word selected was varied from *S* to *S*.

Controls. The controls in this experiment were primarily directed at ascertaining the stimulus for the relevant responses and determining whether the responses, once

obtained, were due to experimental arrangement or to chance. Two steps were taken to permit these determinations. The first of these was the introduction of neutral stimuli at irregular intervals during the training period. Neutral stimuli were also introduced during the test for conditioning and the test for generalization. In this way, responses to undifferentiated stimuli in the total situation could be compared with responses to the conditioned stimulus and to its semantic relative.

The slides used as neutral stimuli were so defined by *E* who carefully selected them to be free of personal reference and obvious emotional coloring. Each neutral word and the word-testing generalization appeared only once for each *S*. The test-list was the same for each *S* within each experimental group but the position of the stimulus testing generalization within the list was allowed to change from *S* to *S*.

The responses of the *Ss* to the neutral stimuli of the test-lists were utilized as controls with which responses to the experimental stimuli were compared. The *Ss* provided their own controls and the controversial problem of combining *GSR*-data was avoided.

Treatment of data. Each *S's* *GSRs* and his reports were examined and treated separately and then were combined. The *GSRs* to each stimulus were measured in millimeters, and the responses to the conditioned stimuli and the stimulus testing generalization were classified as 'significantly' or 'not significantly' larger than the responses to the neutral stimuli of the test-list. In addition to this subjective appraisal of the significance of the experimental responses, a test based upon the standard deviation of the group of seven neutral stimuli was applied.

Fisher's correction was made in the formula for *SD* to compensate for the small samples. A *p* of 0.01 or less was selected as the level of confidence demanded for acceptance of the experimental responses as conditioned or generalized. This objective criterion and the earlier subjective estimate of the significance of the experimental responses were in agreement 82% of the time. This figure reflects 26 errors in judgment. Twenty of the 26 errors occurred when apparent *GSRs* were judged significant at the 1% level when they were actually between the 1% and the 5% levels.

The records, then, were divided into those whose responses to the conditioned stimuli were significant and those whose responses were not. The records of the conditioned *Ss* were again divided into those who made significant responses to the stimuli testing generalization and those who did not.

After this breakdown the *Ss'* post-experimental reports were examined separately and divided with respect to stated expectation of shock upon presentation of the experimental stimuli during the final tests of conditioning and generalization.

Selection of the cases. The need to stop training before complete discrimination between the *CS* and the other stimuli, and the use of only one unreinforced test of conditioning at the end of the experimental run required that the decision to stop training be based upon a judgment. This resulted in frequent, premature termination of training and the rejection of many cases as not conditioned. There were 33 *Ss* from all groups who failed to meet the criterion of conditioning. Of this number 15 *Ss* produced *GSRs* to the *CS* that, although falling short of the criterion, were significant at the 5% level. These cases are not included in the discussion, but they are included in Tables II-IV.

Eight Ss failed to react to the experimental or control stimuli with perceptible GSRs. Instrument failure accounted for the loss of five more Ss. One S complained of discomfort and asked to be excused from the experiment. Thus, 47 Ss were eliminated. In all, 88 persons took part in the experiment. Forty-one of them met the criterion of conditioning.

TABLE II
INCIDENCE OF EXPECTANCY OF SHOCK WITH GSR FOR Ss OF GROUPS I AND II
(X indicates that shock was expected)

Group I							Group II						
S	Conditioned GSR			GSR to test-stimulus			S	Conditioned GSR			GSR to test-stimulus		
	GSR (in mm.)	t	shock ex- pected	GSR (in mm.)	t	shock ex- pected		GSR (in mm.)	t	shock ex- pected	GSR (in mm.)	t	shock ex- pected
1. JW	14	17.18*	X	0	—	—	1. JT	20	4.70*	X	20	2.33†	X
2. BB	13	5.49*	X	0	—	—	2. SB	10	5.58*	X	14	8.23*	X
3. OS	20	3.74*	X	1	—	—	3. PB	30	5.71*	X	27	4.52*	X
4. RS	22	4.19*	X	5	—	—	4. WW	36	9.42*	X	21	5.26*	X
5. CB	25	9.20*	X	21	7.37*	X	5. JHR	42	8.04*	X	37	6.83*	X
6. JZ	28	3.43*	X	12	.33	—	6. TF	20	5.25*	X	6	.49	—
7. RB	11	3.78*	X	0	—	—	7. CP	17	3.60*	X	13	2.39†	X
8. CK	20	6.80*	X	5	—	—	8. RP	7	4.21*	X	0	—	—
9. NK	17	1.98†	X	13	1.30	—	9. HC	18	2.36†	X	13	1.41	X
10. CM	12	3.00†	X	4	.33	X	10. GD	15	2.25†	X	14	2.07†	X
							11. JR	20	3.05†	X	8	—	—
							12. GC	10	1.98†	X	8	.95	—

* Significant at the 1% level. † Significant at the 5% level

TABLE III
INCIDENCE OF EXPECTANCY OF SHOCK WITH GSR FOR Ss OF GROUPS IIIa AND IIIb
(X indicates that shock was expected)

Group IIIa							Group IIIb						
S	Conditioned GSR			GSR to test-stimulus			S	Conditioned GSR			GSR to test-stimulus		
	GSR (in mm.)	t	shock ex- pected	GSR (in mm.)	t	shock ex- pected		GSR (in mm.)	t	shock ex- pected	GSR (in mm.)	t	shock ex- pected
1. BB	20	6.67*	X	4	—	—	1. MS	20	5.80*	X	8	.56	—
2. NS	11	5.90*	X	5	.94	—	2. WW	19	12.48*	X	7	—	—
3. MB	39	45.34*	X	2	—	—	3. RD	24	3.58*	X	0	—	—
4. JD	40	4.46*	X	22	1.49	X	4. GH	22	5.90*	X	4	—	—
5. RS	19	3.24*	X	6	—	—	5. JB	62	24.25*	X	15	5.30*	X
6. JG	16	4.26*	X	10	2.08†	X	6. RH	7	4.08*	X	3	—	—
7. WMcD	41	8.64*	X	2	—	—	7. GP	24	3.27*	X	0	—	—
8. JS	26	4.47*	X	12	.09	—	8. JE	25	3.42*	X	13	1.23	—
9. WH	36	2.86†	X	24	1.26	X	9. JL	50	2.45†	X	39	1.60	X
10. WMc	16	2.10†	X	0	—	—	10. JH	7	2.32†	X	4	.44	—
11. R	18	3.03†	X	4	—	—							

* Significant at the 1% level. † Significant at the 5% level.

RESULTS

Two out of 24 of the conditioned Ss in Groups I (Table II) and IIIa and IIIb (Table III) responded significantly to the stimuli testing generalization. In Group II (Table II), four out of eight of the conditioned Ss responded to the generalization test-stimulus with significantly greater

GSRs. The results for Group IV (Table IV) heavily favored semantic generalization. All but two of the nine conditioned Ss produced significant GSRs upon viewing the generalization stimulus. Thus, there is doubtful

TABLE IV
INCIDENCE OF EXPECTANCY OF SHOCK WITH GSR FOR Ss OF GROUP IV
(X indicates that shock was expected)

S	Conditioned GSR			GSR to test-stimulus		
	GSR (in mm.)	t	shock expected	GSR (in mm.)	t	shock expected
1. RW	27	6.52*	X	17	3.05†	X
2. CH	30	7.91*	X	25	6.46*	X
3. MK	43	4.02*	X	39	3.60*	X
4. JP	47	7.71*	X	40	6.24*	X
5. JK	22	13.20*	X	24	14.47*	X
6. NS	25	3.54*	X	16	1.88	X
7. MG	13	8.70*	X	8	4.35*	X
8. BE	10	3.71*	X	12	4.65*	X
9. TP	6	3.66*	X	7	4.88*	X
10. CC	16	2.29†	X	17	2.59†	X
11. HP	17	2.61†	X	15	2.14†	X
12. CH	12	2.90†	X	9	1.55	X
13. DM	5	2.84†	X	6	3.52*	X

* Significant at the 1% level. † Significant at the 5% level.

evidence of generalization in two of the groups and significant generalization in the other two.

In spite of the inter-group inconsistencies, one relationship, expectation of shock with significantly large GSRs, was found with great dependability. In the cases of the conditioned Ss expectation of shock accompanied *all* of the conditioned GSRs. Out of 41 Ss to whom the stimuli testing generalization were exposed, 19 expected shock and 6 of these failed to produce significant GSRs. There was not a single significant GSR to a test-stimulus that was not accompanied by expectation of shock. If declared expectations had been used to predict the appearance of significant GSRs the generalized responses could have been forecast with 93% success.

DISCUSSION

Three facts emerged from this experiment. Generalization of conditioned GSRs among semantically related materials did not occur with regularity. Generalization occurred more readily in some situations than in others. Expectation of the unconditioned stimulus appeared to be a necessary condition for the arousal of conditioned and generalized GSRs. There were no responses significant at the 1% (or 5%) level that were not accompanied by expectation of the unconditioned stimulus.

The possibility that generalization here resulted from recognition of relationships among the experimental stimuli was suggested in the verbalizations of the Ss. In every case of a significantly large *GSR* to a stimulus testing generalization, S later expressed awareness of the intended relationship between the experimental stimuli, and in every case this knowledge was followed by the inference that shock would follow the appearance of the test-stimulus.

Examples are: *JB* (*S*₅, Group IIIb) who responded significantly to the stimulus testing generalization in Test III(b) made the following post-experimental statement:

BROOK and STREAM are both water. When I saw BROOK I thought I was going to get another shock.

CB (*S*₅ in Group I) reported:

I knew it was about conditioning . . . it was conditioning about word-relationships. When I saw the word it took me by surprise and all I could think of was the picture of the hat. The shock came after the picture so naturally when I thought of the picture I thought the shock would come too.

WMc (*S*₁₀, Group IIIb) who failed to produce a significantly large *GSR* upon seeing the stimulus-testing generalization also failed to see the semantic relationship between the experimental stimuli of Test III(b), and did not hypothesize that shock would follow the test-stimulus. He was able to remember the stimulus used to test generalization but his views as to the purpose of the experiment did not include it. Said he:

Well, every time the word FREEZE came on I got a shock and it seemed that the word FREEZE left an after-image longer than the rest of them. You were trying to see if the electric shock would make me see it better. . . .

Some Ss failed to generalize even though they recognized the relationships between the stimuli. Their statements revealed that they "didn't think the shock would fall on the word as it did on the picture."

Several Ss expected shock to follow the experimental stimuli but nevertheless failed to produce significantly large *GSRs*. These Ss reported that they were not disturbed by the prospect of receiving electric shock. "The shock didn't worry me, it didn't hurt," said one.

The *GSR* is part of the organism's total response to pain and injury. By experience, anticipation of injury comes to evoke the complex of behaviors called emergency reactions and which include the *GSR*. In the expectation of stimuli not grossly painful or which are not threatening or fearful, the organism's defensive responses may not be aroused. If a person can await such stimulation with emotional equanimity he may also remain physiologically passive. Thus the *GSR* is dependent not only upon expectation of a particular stimulus but upon whether the reacting organism views the impending experience with calm or with alarm. In this

study, it appeared that the stimuli which resulted in expectation of shock did not result in GSRs of considerable size unless the electric stimulation was anticipated with fear or apprehension. Expectation of shock was a necessary but not a sufficient condition for arousal of GSRs. The GSRs considered in this experiment followed a sequence of events, including expectation of shock, in which aroused fear of the conditioned stimulus appeared to be a necessary and sufficient condition.

The researches of Cook and Harris which demonstrated arousal of GSRs by instructing Ss that shock would follow a given inadequate stimulus, and Mowrer's study where removal of shock electrodes resulted in failure of arousal of GSRs by a previously, adequate, conditioned stimulus, were consistent with this explanation.²

Situational differentials in the facilitation of semantic generalization of conditioned GSRs. One fact yet to be explained was the frequency with which semantic generalization occurred when several related words (RIVER, CREEK, STREAM) were conditioned and when a fourth related word (BROOK) was presented as the generalization-stimulus. One explanation was that presenting the physically dissimilar but semantically similar words insured discrimination against the physical aspects of the stimuli while facilitating conditioning of an element identical to all of them. Thus, 'wordness' alone, as it occurred with the control stimuli, went many times unreinforced while 'wateriness' was followed by shock every time. It followed then that conditioning in this case may have been to the semantic aspects of the stimuli with resultant evocation of the conditioned response upon presentation of the semantically similar generalization stimulus.

A more parsimonious explanation was suggested by the Ss of Test IV who readily saw the relationships among the conditioned stimuli and who identified the stimuli as 'about water.' The process of hypothesizing the relationships between the conditioned stimuli and the generalization-stimulus with resultant expectation of shock was greatly simplified, being totally revealed in the conditioned stimuli themselves. It was unnecessary, therefore, to formulate further hypotheses during the test-series because the generalization-stimulus fitted into the general hypothesis already generated about the conditioned stimuli and which led to the conclusion of application of shock.

The explanation of the frequency of generalization in Test IV in terms of 'ease of hypothesization' could also be used to explain the relative frequency of generalization in Test II. Ease of hypothesization in this case inadvertently resulted from placing the few pictures, at the disposal of E, in the test-list in order to prevent the test-picture from appearing alone in that list and thereby greatly improving its chances of producing a significant GSR by contrast with neutral words there.

The report of BB revealed the effect of this arrangement.

² S. W. Cook and R. E. Harris, The verbal conditioning of the galvanic skin reflex, *J. exp. Psychol.*, 21, 1937, 202-210. O. H. Mowrer, Preparatory set (expectancy)—a determinant in motivation and learning, *Psychol. Rev.*, 45, 1938, 62-91.

I knew the picture of the hat was going to come sooner or later. I was, more or less, waiting for it. There was the picture of the pipe and the shoe so I knew it was bound to come.

SUMMARY

The results of this experiment revealed that semantic generalization did not occur regularly even in comparable circumstances. Attempts to predict semantic generalization on the basis of accomplishment of conditioning and presentation of semantically related generalization-stimuli resulted in chance success.

All Ss who responded to the generalization-stimuli with significant galvanic skin responses reported (1) awareness of the relationship between experimental stimuli, and (2) expectation of shock. Thus if declared expectancies had been used to predict the appearance of significant GSRs and absence of expectancies to predict non-significant GSRs, the significant responses to experimental stimuli could have been forecast with 93% success.

The reports of Ss who failed to generalize revealed that they also failed to see the relationships between the conditioned and generalization-stimuli, or did not expect shock, or assumed a passive attitude toward the shock.

It was concluded that expectation of shock as a painful or fearful experience was necessary and sufficient to produce responses to the experimental generalization-stimuli in this experiment and that such expectancy was the result of awareness of the existing relationships between the experimental stimuli and experience with the unconditioned stimulus.

PATTERNS OF INTERNAL CONSISTENCY IN MULTIPART SKILLED PERFORMANCES

By EDWARD A. BILODEAU, Tulane University

In research on motor skills, time has been overworked relative to other legitimate variables. Theories about psychomotor skill consist largely in the organization of findings concerning the effects of practice with variable durations of work and rest. The lack of suitable methods of analysis has retarded the study of the internal structure of complex skills. In the field of testing or measurement, however, the central problem has long been the analysis of the internal structure within test-batteries or of the internal pattern of responses within individual tests.

The major difficulty encountered in the area of skills was apparently that of identifying the part-skills, as opposed to the ready identification of test-items in the field of testing. Once identified, there is no reason why part-skills cannot be subjected to the same methods of analysis employed in research on tests. The object of the analysis, too, is much the same: that of understanding or predicting higher-order behavior by means of observation of a number of lesser-order acts combined according to rule.

The following is an example of a psychomotor task where both whole and part activities can be objectively identified (to some degree), and of the research problems that such observations suggest: *S* is required to position hand-controls in response to signals presented by means of display indicators. There may be n indicators to which there correspond n individual controls. An example of a task with $n = 2$ would be two display dials and two hand cranks which position-pointers on Dials 1 and 2, respectively. *S*'s task might be moving the handcranks to compensate for programmed displacements of the pointer.¹

A task such as the one above raises several questions concerning the topic of attention, since *S* must in some way divide his time between Parts 1 and 2. If the two dials are spatially separate, it may be impossible to scan both simultaneously. On the side of response it is obvious that both cranking responses may be simultaneously possible. The degree of simultaneity in response obtained would depend in part on the availability of information regarding required cranking inputs. The proficient *S*, for example, might scan Dial 1 first, and while scanning Dial 2, correct for errors in 1; possibly, correction of error for Dial 1 might be still going

* Received for publication April 9, 1956. This research was carried out at the Skill Components Research Laboratory, Air Force Personnel and Training Research Center, Lackland AFB, in support of Project 7707.

¹ Since we have a typical control system in mind, the programmed displacements will produce signals that are called 'errors.' It should be kept in mind that this error is a function both of the deliberate perturbation of the position of the indicator and of the failure of *S* to compensate properly with the hand crank.

on as corrections for Dial 2 are entered. Another possibility is that both dials are scanned successively, and then corrections are simultaneously entered for 1 and 2. For present purposes, responses on Controls 1 and 2, measured independently, are identified as the parts or components, while simultaneous correction of 1 and 2 is identified as the test-, combination-, or criterion-score.

While there is a possibility of obtaining evidence of substantial time-sharing among two or more responses, it is more likely that over a trial period corrections are both independent and dependent, and that the net amount of sharing beyond chance will not be impressive.² Nevertheless, techniques altering the natural manner of scanning have been evaluated and have shown the possibility that increases in time-sharing are possible, though not necessarily by alteration of the statistical structure of the events within the trial.³

The papers mentioned have dealt with the practice of tasks of constant n . No comparison of the relationships between part-tasks and indices of simultaneity have yet been made *between* tasks which differ in the number of components presented. The present experiment attempts to do this by systematically varying the number of components defining the task. Measures of response-proficiency for each of the n components are taken, as are measures of the degree to which two or more components are associated in time.

The experiment is intended to bear on a rather large area concerning the nature of the composition of skill, but not necessarily to provide definitive answers. Typical questions about the area of how skills are composed are: (1) How is proficiency on the j th component related to the total number n of components presented? (2) What is the relationship of the proficiency exhibited on a task of n components between combination scores of successive orders, i.e. ab , abc , $abcd$. . . ? (3) What is the correlation between components, between one component and the combination score for a task of n components? What is the pattern of multiple correlations?

The present experiment provides data for the above questions and provides alternative ways for examining the relationships between the parts and combinations of parts of a psychomotor tracking task.

METHOD

Apparatus. One unit of the SAM Multidimensional Pursuit Test (CM813E) was used.⁴ The standard task requires S to center four meter-pointers in a compensatory manner by appropriate manipulation of four controls. The display panel is shown

² E. A. Bilodeau and I. McD. Bilodeau, The contribution of component activities to the total psychomotor task, *J. exp. Psychol.*, 47, 1954, 37-46.

³ E. A. Bilodeau, Variations in knowledge of component performance and its effects upon part-part and part-whole relations, *ibid.*, 50, 1955, 215-224.

⁴ A. W. Melton (ed.), *Apparatus Tests: AAF Aviation Psychology Program Research Report No. 4, Restricted Supplement (declassified)*, 1947.

in Fig. 1 and, for simplicity, the four meters and their corresponding controls are called components S, H, A, and B. Both A and B represent movements of a stick-control held in the right hand; A is associated with forward and backward motions, and B is associated with left-right motions. H represents the 'rudder' pedals, and S is associated with a 'throttle' or small lever held in the left hand. Time-on-target was accumulated whenever any one of the meter-pointers was held within one scale division of center, plus or minus. Measures were also taken of the time two or

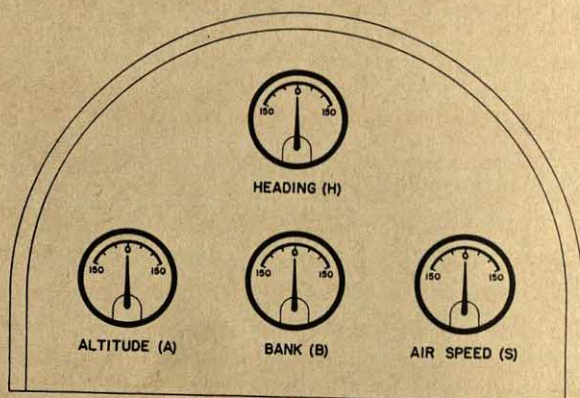


FIG. 1. MODIFIED DISPLAY PANEL OF THE MULTIDIMENSIONAL PURSUIT-TEST

more pointers were within the scoring bands simultaneously, *i.e.* AB, HAB, and SHAB.⁵ These are combination scores or measures of time-sharing between components; SHAB, for example, represents the time-on-target score for correct performance of all four components simultaneously. Cardboard covers were provided for any meter not in use.

Procedure. An experienced man was used as S. This S had received a very considerable amount of prior practice and had demonstrated a stable level of performance. The task was explicitly defined as keeping centered all uncovered meter-pointers, all of the time. It was made clear that this was not to be interpreted as keeping the pointers within the scoring band, but that the task consisted in keeping all pointers on center.

Four different tasks were administered. Task A consisted in practicing A; Task AB in practicing A and B; Task HAB in practicing H, A, and B; and Task SHAB in practicing S, H, A, and B. The component or components not in use were covered by the cardboard covers, and S was instructed to attend only to the uncovered displays. A standard method of scanning the meters was used, *i.e.* H to A to B to S ($n = 4$), or H to A to B ($n = 3$), or A to B ($n = 2$). If a control was not operative, S held it in the neutral position.

⁵ The notation may cause some difficulty. In Task S, H, A, B, for example, four parts are presented. AB represents a response-measure, $A \times B$ represents the product of two part-response measures and A, B etc., not preceded by the word *Task*, represents the statistical treatment of Parts A and B in a correlational sense.

Table I summarizes the parts defining the four tasks as well as the measures of response appropriate to each. In general, A, B, S and H are to be considered as the parts of the tasks, and AB, HAB and SHAB as the criterion combinations. It should be noted that parts A and B are common to all criteria, that AB is included within HAB and SHAB, and that HAB is included within SHAB. Response measure SHAB is considered a higher-order criterion than HAB, and HAB is considered an order higher than AB.⁶

All trials were 60 sec. in duration, and rests between successive trials were 120 sec. Forty-eight trials were administered on each of 4 days, 24 trials in the morning,

TABLE I
FOUR TASKS AND THEIR MEASURES OF RESPONSE

Number of components	Task practiced	Measure of response
1	A	A
2	AB	A, B, AB
3	ABH	A, B, H, AB, HAB
4	ABHS	A, B, H, S, AB, HAB, SHAB

24 in the afternoon. Within a half day, 30-min. rests were interpolated after every eight trials.

The cycling mechanism of the apparatus administered eight trials automatically. Each of these eight trials differed from the others in some minor respects. To control this potential source of variability the following order of practice was administered. Each task appeared twice in any one set of eight trials. By the end of four days, each task had been paired six times with any given trial. As a consequence there were 48 estimates of performance for each response-measure of each task.

RESULTS

Though the parts were not equally difficult, they did not vary greatly in the distribution of scores which they produced. For Task AB the parts A and B were not significantly different at the 5% level, while for Tasks HAB and SHAB, complete interpart comparisons gave *F*s significant beyond the 1% level. The greatest difference between mean part-scores was A of 0.74 and B of 0.79, A of 0.62 and B of 0.71 for Tasks HAB, and SHAB, respectively. Evidently, the relative difficulties of two parts varies with the number of parts in the task.

The means and *SD*s ($N = 48$) of the more important time-on-target proportions are presented in Table II. The table shows that as the number of components is increased time-on-target for one measure of response decreases (see columns A, AB, and HAB). No systematic trends in the *SD*s appear either within or between tasks. The obtained means and *SD*s

⁶ See Footnote 5.

of Table II provide estimates of trends and magnitudes of scores to be expected as components are successively added. The data should be inspected with n as the parameter, and then again with the measure of response as the parameter.

Estimates of the obtained values of Table II are presented under the columns labeled $A \times B$, $H \times A \times B$, and $S \times H \times A \times B$. To obtain

TABLE II
OBTAINED AND PREDICTED PROPORTIONS OF TIME-ON-TARGET FOR FOUR TASKS AND
FOUR MEASURES OF RESPONSE

Task		A	AB	A × B	HAB	H × A × B	SHAB	S × H × A × B
		(obtained)	(obtained)	(predicted)	(obtained)	(predicted)	(obtained)	(predicted)
A	mean	.973						
	SD	.009						
AB	mean	.894	.795	.788				
	SD	.034	.066	.062				
	r^2			.97				
HAB	mean	.730	.592	.588	.448	.450		
	SD	.079	.092	.088	.101	.088		
	r^2			.94		.92		
SHAB	mean	.623	.457	.442	.288	.295	.199	.195
	SD	.076	.074	.073	.089	.074	.077	.060
	r^2			.80		.83		.78

these estimates it was assumed that any combination score C represented the probability of the joint occurrence of n independent events. In terms of the measures, then,

$$A_o \times B_o \times \dots = C_n, \quad [1]$$

where subscript o identifies the obtained proportions of time-on-target. The utility of Equation [1] has already been verified in previous papers.⁷ In Table II under Task AB, Criterion AB is 0.795 and $A \times B$ is 0.788 or a slight underestimate of the obtained value. Whether the task consisted of 2, 3, or 4 parts, Equation [1] gave values numerically similar to the observed values. That the relationship is fairly strong is shown by the squares of r . The correlations were obtained by pairing the arrays of 48 obtained and predicted scores. No value in Table II is less than 0.78, and the patterning of the squares of r suggests that the relationships are weaker as a function of increasing number of components, though there is no trend as a function of the complexity of criterion for a given task.

Variance of criterion. If the combination-score represents the joint probability of being on target, as expressed by Equation [1], then the log of the combination-score will be linearly related to the logs of its parts, as shown in Equation [1a]. It is assumed that all parts are presented.

⁷ Bilodeau, *op. cit.*, 215-224. Bilodeau and Bilodeau, *op. cit.*, 37-46.

$$\log C_n = \log A + \log B + \dots \log n \quad [1a]$$

The correlation, then, of log SHAB with log A, for example, would be a satisfactory way of estimating the relationship between a criterion combination and one of its parts. This procedure gave satisfactory empirical rectification for plots of part-scores against criterion-scores and for plots of the part-scores against one another. The following is a detailed account of within-task consistencies based upon standard correlational procedures and the logs of the measures of response.³

The major purpose of this analysis was to determine the amount of common variance between part-predictors and the combinations of scores used as criteria and is an illustration of the item-test approach. The relationships were expected to be close, of course, because scoring in a combination implies scoring in the parts, and because of similar findings in preceding work with the multiplicative rule. The analyses were accomplished by means of bivariate and multiple correlations, Task SHAB receiving the greatest attention because it was the task with the

TABLE III

INTERCORRELATIONS AMONG SEVEN MEASURES OF RESPONSE IN TASK S,H,A,B
(Correlations were computed using the log transformation of original scores.)

Measures of response	Measures of response					
	A	B	H	S	AB	HAB
B	-.04*					
H	.27	.32				
S	.20	-.10	.26			
AB	.73	.46	.43	.19		
HAB	.58	.48	.77	.25	.84	
SHAB	.51	.34	.72	.53	.70	.88

*With $df = 46$, an r of 0.37 and 0.29 is required to reach the 1% and 5% levels of significance, respectively.

greatest number of part-tasks. The intercorrelations among the logs of the response-measures of Task SHAB are shown in Table III. The intercorrelations *among the parts* are low, as expected, though the value for B,H (0.32) is higher than desired. In general, the correlation of a part-score with a combination not including that part (irrelevant part) leads to a value less than that obtained by correlating a part-score contained within the combination (relevant part).

The squared correlations of parts against combination-scores are presented in Table IV to show the pattern of common variances. Part A, for example, has 54, 33 and 26% common variance with the combination scores AB, HAB, and SHAB, respectively. Part S does not correlate highly with AB or HAB, but does account for 28% of the variance in SHAB. The bottom three rows of Table IV show the advantage of using two, three, or four parts in predicting the combinations, using the coefficient of multiple correlation, R . Adding a relevant part appears to increase the measure of the relationship while adding an irrelevant part appears to have no effect. Though the use of all relevant parts accounts for a sizable amount of vari-

³ Much of the correlational work reported in this paper was accomplished by IBM methods under the direction of Dr. Joe H. Ward, Jr., of the Personnel Research Laboratory, Air Force Personnel and Training Research Center.

ance, about 20% of the variance in the criterion remains unexplained. As the number of relevant parts is increased, greater amounts of variance are accounted for, yet there is a diminishing return. It is also true that a constant number of predictors does more poorly the higher the order of the criterion, and that when all parts are used the value of R^2 remains fairly independent of the number of task elements.

More common variance than shown in Table IV can be demonstrated by using

TABLE IV

PROPORTION OF CRITERION-VARIANCE ACCOUNTED FOR BY PREDICTION FROM PARTS
IN TASK SHAB

(Values are squared coefficients of correlation on basis of log measures of response.)

Predictors	Criterion		
	AB	HAB	SHAB
A	.54	.33	.26
B	.21	.23	.12
H	.18	.59	.52
S	.03	.06	.28
AB	.78	.59	.39
HAB	.79	.83	.65
SHAB	.79	.83	.77

TABLE V

PROPORTION OF VARIANCE IN CRITERION ACCOUNTED FOR BY PREDICTIONS
FROM PARTS AND COMBINATIONS OF PARTS IN TASK SHAB

Predictors	HAB Criterion	Predictors	SHAB Criterion
AB	.70	HAB	.78
AB,A	.70	HAB,A	.78
AB,B	.71	HAB,B	.79
AB,H	.90	HAB,H	.79
AB,S	.71	HAB,S	.89
		HAB,AB	.79

a low-order combination and a relevant single part in a multiple correlation against a higher-order criterion. This analysis is shown in Table V where, for example, R^2 for AB and H against HAB is 0.90, R^2 without H is 0.70, and R^2 for AB and an irrelevant or redundant (*i.e.* a part entering into the lower-order criterion) part against HAB is no greater than 0.71. When SHAB is used as the criterion the values of R^2 are generally larger than with HAB as a criterion, but both are predicted about equally well by the critical variables. When SHAB is the criterion the critical comparison is with HAB and S as the predictors, the other predictors shown in the table in a sense serving as control comparisons.

Another analysis is presented in Table VI where the question of determining the highest obtained percentage of common variance has been asked. Here, the seven response-measures have been arranged as seven criteria along with a number code. The predictor variables are shown in the first column and are identified according to the code given in the second column. The values of R^2 in the third column are the highest yet obtained. In parentheses, however, are the values of R^2 when only the relevant parts are used, showing substantial, though lesser, relations.

As noted before, the values within parentheses can be raised still further by considering lower-order criteria as predictors.

To this point, two procedures have been used to account for the variance in the criterion-scores. One method is based upon the multiplicative relation between parts, and the other treats the simultaneous regressions of part performance upon the criterion. If we consider AB as the criterion for the moment, and use C as the notation for the response measure AB, the correlations can be expressed as $r_{C \cdot A \times B}$ and $R_{C \cdot A, B}$. The procedures are compared in Table VII for each of three tasks. The

TABLE VI
PROPORTION OF VARIANCE IN CRITERION ACCOUNTED FOR BY PREDICTIONS
FROM PARTS AND COMBINATIONS OF PARTS IN TASK S,H,A,B, BASED
ON MULTIPLE CORRELATIONS PREDICTING ONE SCORE FROM THE
REMAINING MEASURES

Predictors	Code & Criterion Predicted		R ²
2, 3, 4, 5, 6, 7	1	A	.72
1, 3, 4, 5, 6, 7	2	B	.57
1, 2, 4, 5, 6, 7	3	H	.74
1, 2, 3, 5, 6, 7	4	S	.54
1, 2, 3, 4, 6, 7	5	AB	.88 (.78)
1, 2, 3, 4, 5, 7	6	HAB	.95 (.83)
1, 2, 3, 4, 5, 6	7	SHAB	.89 (.77)

TABLE VII
A COMPARISON OF TWO PROCEDURES ACCOUNTING FOR VARIANCE IN MEASURE OF
RESPONSE AB FOR EACH OF THREE TASKS

Method of Prediction	Task		
	AB	HAB	SHAB
AB · A,B	.97	.95	.78
AB · A × B	.97	.94	.80

entries give no reason to believe that the different procedures give different results. The reason $r_{C \cdot A \times B}$ and $R_{C \cdot A, B}$ give similar results is fairly clear: A and B can be substituted for AxB since both parts are fairly highly correlated with AxB and yet are uncorrelated with each other.

DISCUSSION

The design of this study was specifically directed toward an analysis of the events within tasks ranging from one to four parts. The analysis was always concerned with the relations existing at S's level, never at the level between Ss. Still, more than one S might have been usefully employed to obtain better estimates of within-S events. Since the data should be collected from Ss who are well experienced at performing the assigned tasks, the collection of data for a number of Ss would be time consuming and

costly. A full-scale effort should only be undertaken on the basis of evidence that the effort would be remunerative. The high degrees of consistency in the present data make generalizations about the single *S* fairly safe, and suggest a number of promising results to be obtained from more costly investigations.

The findings bear upon two general topics. A better idea of the functional relationship between performance on a part and the number of task components has been provided. Knowledge of this relationship is essential to describing principles of task-difficulty, and to an intelligent estimate of the level of performance that can be expected of proposed psychomotor devices, *i.e.* what would happen to proficiency were a part added to or subtracted from the device. More importantly, however, the present results bear upon the rules of composition of an aggregate of individual parts of a task, or upon the relations between the constituent responses of the whole, and of the relation between the constituent responses and various orders of criteria. Two types of analyses were used. In the first, the multiplicative rule of independent probabilities was used to estimate the probability of the joint occurrence of scoring in two or more parts, and these estimates were compared with the observed values by means of correlational procedures. The second procedure was based upon the more orthodox method of estimating the regression of part scores upon the total score. The correlations based upon comparisons such as $C \cdot A \times B$ and $C \cdot A, B$ were impressively large, and nearly equivalent.

Fairly extensive analyses were made of the contribution in variance of part-scores to the total test-scores. It was shown that part-scores tend to correlate weakly, if at all, with each other, but that each contributes a considerable, though varying, amount to the variance in total-scores. The results showed too that a score based on the proficiency on a part not included within the total task is fairly useless as a predictor.

The methods and findings of this study should prove useful in constructing a theoretical model for complex psychomotor skills, though the model would have obvious limitations. Two limitations stem directly from what has been considered the basic datum. All scores have been time-on-target and so do not include any index or other measure of responses not on the target; all scores represent a cumulation of discrete events within a trial (1 min.) and so do not identify more molecular processes known to exist. Thus, the analysis of time-on-target scores should be understood as a relatively molar approach to the *S-R* fractionation of time-sharing activities.

The importance of prediction for tasks of *n* parts on the basis of ob-

servation of the practice of $n-k$ parts is well recognized. The analysis would have to be based on empirical relations determined from longitudinal sections of behavior. While the present study contains data suitable for such analyses it has received no great stress here. One method of estimating the combination-scores is to compute the regression of A upon the number of tasks practiced, and then to assume (primarily for simplicity) that all parts are of equal magnitude. Accordingly, by the method of least squares, a loss of 0.120 points was estimated for the addition of successive parts. If the proportion of time-on-target for A for Task A is 0.973, then for Task AB, the estimated time-on-target proportion for any one component becomes $0.973 - 0.120$ or 0.853. The combination-score AB can then be estimated by the square of 0.853 or 0.728. An expression of the above technique is given in Equation [2],

$$C_n = [A_1 - .120(n-1)]^n \quad [2]$$

where C_n represents the criterion combination, A_1 the response measure for Task A, and n the number of parts in the task as presented to S.

CONTRAST, ASSIMILATION, AND THE EFFECT OF CENTRAL TENDENCY

By WILLIAM A. WATSON, The Rand Corporation

Previous studies in psychophysics have shown that when a subject (*S*) makes judgments about a series of stimuli, which vary in some dimension, his judgments of any one stimulus is affected by the other members of the series. Specifically, when two physically equal stimuli are presented successively at short time-intervals, the judgment of the second in terms of the first depends upon the position of these stimuli in relation to the series. For example, when the stimuli are from the large end of the series, *S* exhibits a tendency to judge the second as larger than the first; when they are from the small end, he tends to judge the second as smaller than the first. The Time-Error (*TE*) for large stimuli tends to be negative; for small stimuli, to be positive.

Hollingworth suggested that this phenomenon be called the "Central-Tendency Effect."¹ Although a number of theories have been proposed to account for this result, they reduce to two: the Assimilative Theory² and the Contrast Theory.³ Both theories hypothesize a central level for the series of stimuli as a whole, similar to Helson's "Adaptation Level." Stimuli above this level are affected in a way which is just the reverse of stimuli which lie below it. According to the Assimilative Theory, the traces of the standard stimuli assimilate toward the level during the intervening time-interval, hence, when a comparison-stimulus equal in size to the standard is finally presented, it appears more extreme than the standard. According to the Contrast Theory, every stimulus is contrasted with the central level and it thus appears to be more extreme than it should, but this exaggeration lasts only for the actual duration of the stimulus itself.

With these basic assumptions, the two theories both predict positive time-errors for stimuli below the central level and negative time-errors

* Received for publication September 14, 1956. This paper was adapted from a doctoral dissertation submitted to the University of Michigan. The author is indebted to Dr. A. B. Kristofferson for his assistance on this study.

¹ H. L. Hollingworth, The inaccuracy of movement, *Arch. Psychol.*, 2, 1909 (No. 13), 1-87.

² Otto Lauenstein, Ansatz zu einer physiologischen Theorie des Vergleichs und der Zeitfehler, *Psychol. Forsch.*, 17, 1933, 130-177.

³ C. C. Pratt (Time-errors in the method of single stimuli, *J. exp. Psychol.*, 16, 1933, 798-814) indicated how the Contrast Principle could be used to account for Lauenstein's results.

for stimuli above this level. In the case of absolute judgments both theories predict that stimuli will be categorized as being more extreme than they deserve. The explanation of this is obvious in the case of the Contrast Theory. In the case of the Assimilative Theory it is assumed that the traces of previous stimuli, upon which *S*'s subjective categories are based, have assimilated toward the level thus compressing the scale of categories relative to the physical scale. When a 'fresh' stimulus is presented for classification it then appears more extreme than it should.

OBJECT

The present experiments were designed to obtain evidence relevant to the two theories in question. The general plan was to set up two psychologically distinct classes of stimuli on a single dimension by differentiating them on a number of characteristics in addition to the attribute for which the central tendency was to be measured, *i.e.* size. In this way it was hoped that two central tendencies could be established, one for each class of stimuli. These two classes were of different ranges hence the hypothetical level for one group would be larger than that for the other. Once the classes were established it would be possible to present the critical test-stimuli. This presentation would consist of two simultaneously presented stimuli, equal in size (and lying between the two levels), and each representing one of the original classes. Under these circumstances the Contrast Theory would predict that each test-stimulus would be contrasted with its own level and that the stimulus associated with the smaller class would be judged larger than the stimulus associated with the larger class.

The Assimilative Theory, on the other hand, predicts either of two alternatives depending upon what assumption is made concerning how quickly the process of assimilation starts. If there is some assimilation immediately, then the test-stimulus associated with the larger class should be judged larger. If it takes time for assimilation to occur then the test-stimuli should look equal.

METHOD

Subjects. A total of 164 *Ss* (men and women) were used: 47 in Experiment I, 44 in Experiment II, and 73 in control experiments. All were students, mostly undergraduates, at the University of Michigan.

Apparatus. The stimuli consisted of circular areas of light projected from approximately 14 ft. on a 4 × 8 ft. white screen. The sizes of the standard stimuli used in Experiments I and II are shown in Table I. The only difference between these two experiments lay in the sizes of the stimuli used. In Experiment I the

ranges of the two classes did not overlap, in Experiment II they did. The sizes of the stimuli used in the Control Experiment were the same as those of Experiment I. In the latter case, however, the two classes were not used together. The small stimuli were presented to one group of Ss (Group A, $N = 37$), while the large stimuli were presented to a separate group of Ss (Group B, $N = 36$). Two comparison-stimuli were used with each standard, one larger than the standard by about 2% and the other smaller than the standard by about 4%. The circles were projected from film strips, a different frame being used for each standard. The sizes of the comparison-stimuli were obtained by moving the projector forward and backward

TABLE I
SIZES OF STANDARD STIMULI IN INCHES

Experiments	Small class				Large class			
	3.13	3.74	4.57	5.50	18.38	20.02	21.79	23.86
I and Control	6.83	8.04	9.48	11.12	9.72	11.40	13.38	15.72
II	10.31							
Test-stimuli								

on a set of rails from the standard position. Changes in focus over the range used were indiscriminable; changes in brightness were compensated for by the insertion of resistors into the circuit. The reflected brightness of the circles was approximately 1.3 foot-lamberts. The room itself was in semi-darkness, reflecting approximately 0.3 foot-lamberts from the screen. A timing mechanism presented a stimulus for 1 sec. and then, 5 sec. later, the second stimulus for 1 sec. The time-interval between successive comparisons was varied between 5-10 sec. Two projectors were used in Experiments I and II, mounted one above the other. S's eye level was stabilized at about 45 in. above the floor with a chin-rest, and the two projectors were so arranged as to project their respective circles equal distances (about 15 in.) above and below S's eye-level. The single projector used in the Control Experiment projected circles at approximately S's eye-level. Ss wore cylindrically-shaped cardboard masks which restricted their field of vision somewhat to prevent distractions.

Procedure. Single experimental sessions involved groups ranging in size from 1 to 10 Ss, who recorded their own responses. The stimuli presented in each experiment were arranged in two parts: a preliminary series and a test-series. The first and last portions of every preliminary series were counterbalanced as thoroughly as conditions permitted. The preliminary series were divided into a number of cycles—a cycle consisting of the random presentation of the four standard stimuli in a class along with randomly determined comparison-stimuli. During the preliminary series, S's task was always to compare the first (standard) stimulus with the second (comparison) stimulus and to judge which of the two was larger. 'Equal' judgments were not permitted. Throughout the course of Experiments I and II the top and bottom projectors were used alternately; one cycle was presented on the top, the next cycle on the bottom, etc. Top and bottom circles were not only distinguished by their position (top or bottom) but also by their colors (green or yellow) and sizes (large or small). Filters were used to produce colored stimuli, and Eastman Wratten No. 102 for green circles and Eastman Wratten No. 15 with a neutral density filter for the yellow circles. For any S then, there were four standard

circles of a particular range of sizes and a certain color located on top and another four standards with a different range of sizes and a different color located on the bottom. The preliminary series consisted of 12 top cycles alternating with 12 bottom cycles for a total of 96 trials. On Trials 97, 106, and 115, test-stimuli were presented. These consisted of a simultaneous presentation of two circles, equal in size, one on the top, the other on the bottom, and each with its appropriate color. The trials between test-stimuli were composed of alternating cycles like those in the preliminary series.

In both experiments there were four experimental groups, each with a different

TABLE II
COMPOSITION OF EXPERIMENTAL GROUPS IN EXPERIMENTS I AND II

Group	Top stimuli		Bottom stimuli	
	color	size	color	size
I	yellow	small	green	large
II	green	small	yellow	large
III	yellow	large	green	small
IV	green	large	yellow	small

combination of colors and sizes in the two positions. The composition of the four groups is shown in Table II.

The Control Experiment differed from Experiments I and II in that only 20 cycles were presented in the preliminary series and all stimuli were achromatic. Judgments of the test-stimuli for the Control Experiment will not be discussed.

RESULTS AND DISCUSSION

Preliminary series. Before any computations were made, the judgments for all Ss were combined within each of the experiments. Points of Subjective Equality (PSE) based on medians derived from linear interpolations and time-errors (TE) were computed on the grouped data for the preliminary series judgments in each experiment. The time-errors were expressed as percentages of the standard stimuli. Tables III, IV, and V give the results of this analysis while the same is shown graphically in Fig. 1. In general it can be seen that as the stimuli in any group of stimuli get larger, the time-errors become more negative. This is characteristic of the effect of central tendency. The presence of an over-all negative time-error is also apparent from the fact that most of the time-errors are negative.

The only question with respect to judgments in the preliminary-series to be considered here is whether two central tendencies were established within each of the experiments. Ideally, two central tendencies within an experiment would produce a distribution of time-errors which was positive for small stimuli gradually changing to negative for large stimuli

within each class. The results of Experiment II approximate this fairly well except for the confounding effect of the over-all negative time-error. At any rate the two smallest stimuli in the large class are more positive than the two largest stimuli of the small class. This suggests that the

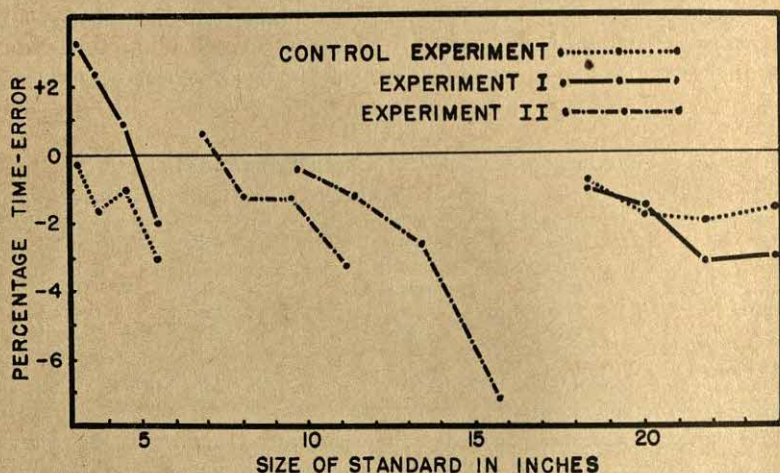


FIG. 1. PERCENTAGE TIME-ERRORS FOR JUDGMENTS IN THE PRELIMINARY SERIES OF EXPERIMENTS I, II, AND CONTROL EXPERIMENT

TABLE III
JUDGMENTS IN THE PRELIMINARY SERIES OF EXPERIMENT I
(Each PSE based on 563 judgments)

	Small class				Large class			
Size (standard)	3.13	3.74	4.57	5.50	18.38	20.02	21.79	23.86
PSE	3.23*	3.83*	4.61	5.39	18.19	19.70	21.10	23.13
% TE	+3.19	+2.41	+ .86	-2.00	-1.03	-1.60	-3.17	-3.06

* Based on linear extrapolation.

stimuli in the two classes were being related psychologically to two different levels.

The results of Experiment I are more ambiguous. In this case it could be argued that the distribution of the time-errors for the two classes suggests a single central tendency, rather than two, which could have arisen if all the stimuli involved in the experiment were related to a single level. Against this reasoning is the steepness of the negative slope for the small class and the fact that the time-error for the largest stimulus in the small class is more negative than the two smallest stimuli in the large group,

and this despite the fact of the great difference in size between the two groups.

If two levels were present then there was at least some interaction between them. If two completely independent psychological classes were present in Experiment I, then the time-errors should have looked like those of the Control Experiment in which the two classes were actually presented separately with different Ss. It is apparent that the time-errors for the small class became more positive and for the large class more negative when they were used together in the same experiment—the effect

TABLE IV
JUDGMENTS IN THE PRELIMINARY SERIES OF EXPERIMENT II
(Each PSE based on 528 judgments)

	Small class				Large class			
Size (standard)	6.83	8.04	9.48	11.12	9.72	11.40	13.38	15.72
PSE	6.87	7.94	9.35	10.76	9.67	11.26	13.02	14.59*
% TE	+ .59	-1.24	-1.37	-3.24	-.51	-1.23	-2.69	-7.19

* Based on linear extrapolation.

TABLE V
JUDGMENTS IN THE PRELIMINARY SERIES OF THE CONTROL EXPERIMENT

	Group A*				Group B†			
Size (standard)	3.13	3.74	4.57	5.50	18.38	20.02	21.79	23.86
PSE	3.12	3.68	4.52	5.33	18.23	19.66	21.35	23.47
% TE	-.32	-1.60	-1.09	-3.09	-.82	-1.80	-2.02	-1.64

* Each PSE based on 740 judgments.

† Each PSE based on 720 judgments.

being greater the more extreme the stimuli. These results could be accounted for on the assumption that two levels were involved and that they were drawn together on the basis of the attribute of size. Such a mutual assimilation would account for the fact that in both experiments the stimuli in the large class were more negative in general than the stimuli in the small class. It is this kind of reasoning which best accounts for the data from the preliminary series since, as will be pointed out, the judgments of the test-stimuli cannot be accounted for on the assumption that there is only one level in operation.

Test-stimuli. Summaries of the judgments of the test-stimuli for the four experimental groups of Experiments I and II are given in Table VI. In both experiments a contrast-effect is indicated since in both experiments

the stimulus judged larger tended to be the one associated with the small class of stimuli.⁴ Chi-squared tests were run on the combined groups within each experiment (indicated by the totals in Table III). In both cases the chi-squared was significant at beyond the 1/10% level of significance (Experiment I, $X^2 = 16.70$, $df. = 1$; Experiment II, $X^2 = 59.42$, $df. = 1$). In addition to the immediate question of a contrast-effect, the judgments of the test-stimuli also indicate that the data of the present experiments cannot be accounted for on the assumption of there being

TABLE VI
JUDGMENTS OF TEST-STIMULI IN EXPERIMENTS I AND II
Test-stimulus judged larger

Group	Experiment I			Experiment II		
	top	bottom	totals	top	bottom	totals
I	35	4	39	21	12	33
II	25	5	30	19	11	30
Totals I & II	60	9		40	23	
III	10	20	30	8	22	30
IV	6	36	42	11	27	38
Totals III & IV	16	56		19	49	
Grand total	76	65	141*	59	72	131†

* Three judgments from each of 47 Ss.

† Three judgments from each of 44 Ss; one S failed to make one judgment.

just a single level in operation for the series as a whole. If there were only one level, then any influence exerted by it on the test-stimuli would affect them both in exactly the same way. In such a case, the Ss would have no preference for either of the test-stimuli.

Theoretical compromise. The results of the present experiments support the hypothesis that there is a contrast-effect in operation in the central tendency—a result which the Assimilative Theory cannot explain. There is also a suggestion in the data that the amount of contrast is greater the farther the stimulus deviates from the level. About 82% of the judgments of the test-stimuli were consistent with the contrast principle in Experiment II where the test-stimuli were both remote from their respective groups,

⁴Results similar to these were reported by Brunswik and Herma (Probability learning of perceptual cues, *J. exp. Psychol.*, 41, 1951, 281-290) using lifted-weights. Likewise, some of the results of figural after-effect experiments bear a close resemblance. It should be kept in mind, however, that in the present study Ss were free to fixate on each of the stimuli of the preliminary series in turn with the probable effect that most of these presentations appeared in roughly the same retinal location.

whereas only about 68% of the judgments were consistent with the contrast principle in Experiment I where the test-stimuli deviated less. Actually, the Contrast Theory of the central tendency must postulate such a relation if it is to account for the increase in the size of the time-error as the stimuli get more extreme. In a like manner, the Assimilative Theory would have to postulate greater assimilation for more remote stimuli.

The results of the present experiments do not, however, completely decide the issue for the Contrast Theory. The Contrast Theory has the decided disadvantage of not being able to handle comfortably the data of previous experimenters on the effect of the standard-comparison time-interval on the central tendency. It has been found that as this time-interval is increased, the central tendency becomes more exaggerated (the slopes of the curves as in Fig. 1 become steeper).⁵ The Assimilative Theory would account for this simply: the longer the time-interval between the standard and comparison-stimulus, the greater the assimilation of the trace of the standard and thus the greater the time-error. For the Contrast Theory to hold, the contrast exhibited by the comparison-stimuli would have to increase by larger amounts as this time-interval increases. There is no basis for such a notion. Besides, a more reasonable alternative is available.

The fact that a contrast effect was demonstrated in the present experiments does not eliminate the possibility of an assimilative effect as well. Thus it would be possible for the *psychological* stimulus-value of a stimulus to be displaced immediately away from the level (contrast) and then, over a period of time, to revert back toward the level (assimilation).⁶ If the assumption is made that more remote stimuli are successively contrasted and assimilated to a greater extent than stimuli near the level, then most of the data can be accounted for, including the effect of the standard-comparison time-interval on the time-errors. There is no evidence available, however, which would indicate whether the assimilated trace ever gets back to (or perhaps goes beyond) the 'original' stimulus-value of the stimulus; that is, its stimulus-value out of context.⁷

This last statement brings into focus the question of the essential mean-

⁵ J. G. Needham, The effect of the time-interval upon the time-error at different intensity levels, *J. exp. Psychol.*, 18, 1935, 530-543.

⁶ Theodore Koester and W. N. Schoenfeld (The effect of context upon judgments of pitch differences, *J. exp. Psychol.*, 36, 1946, 417-430) first suggested the possibility of both contrast and assimilation operating successively (on the standard stimulus only), although they were trying to account for somewhat different data.

⁷ The author also has some evidence that for stimuli which deviate from the expected range by a considerable amount the contrast effect is retained, but the assimilation gradient is not nearly so steep, and indeed may not be present at all.

ing of the concept of contrast. When the statement is made that the stimulus-value of a stimulus is displaced in a direction away from the level, this implies a displacement away from some value which the stimulus would have exhibited were it not in a context, its 'original' value. In the present experiments, it will be noted, the stimulus-values of stimuli *in* and *out* of context were never compared. What were compared were the relative values of two stimuli (the test-stimuli) each in a different context. Their *out* of context values were assumed to be the same since they were physically the same size and since color and position were found to be unimportant factors in determining these judgments. Since actual judgments indicated that these stimuli were not seen as equal, it was inferred that the context had altered the stimulus-values of each.

The objection might be raised that the contrast-effect noted in the judgments of the test-stimuli could have been due to the interaction of the two stimuli and that such an effect might not be present with the presentation of a single stimulus. Although this is a theoretical possibility it then becomes difficult to see how the presence of a contrast effect for a single stimulus could ever be detected empirically.

The theory outlined above leads to at least one testable hypothesis. If, in the setup described for Experiments I and II, a time-interval is introduced between the two test-stimuli, the amount of contrast should be reduced.

SUMMARY

The main purpose of the present experiments was to obtain evidence relevant to the Contrast and Assimilative Theories of the Central Tendency. Ss made size-judgments on two classes of circular stimuli using the method of constant stimuli. Classes were differentiated with respect to color, position, and size, the expectation being that under these circumstances two rather than one Central Tendency would be established. Presentations of the test-stimuli following familiarization consisted of the simultaneous presentation of two stimuli, equal in size, and each representing one of the original classes.

The distributions of time-errors were most readily interpreted as representative of two separate but interacting cognitive classes. The existence of a contrast-effect in the judgments of the test-stimuli was indicated in that stimuli representative of the smaller of the two classes tended to be judged larger than stimuli representative of the larger class. A compromise theory was suggested, involving contrast followed by assimilation.

DISCRIMINATION-REVERSAL IN THE FISH

By JEROME WODINSKY, American Museum of National History, and
M. E. BITTERMAN, Bryn Mawr College

The experiment to be reported here is concerned with the course of discrimination-reversal in the fish (*Tilapia macrocephala*) under conditions analogous to those employed for the study of reversal-learning in higher animal forms. The Ss were trained in a two-choice situation to strike a horizontally or a vertically striped target (irrespective of lateral arrangement, which varied from trial to trial), and the established preference was then reversed repeatedly. For example, Ss trained originally to horizontal stripes were retrained to vertical stripes, then to horizontal stripes, and so forth until as many as 15 reversals were accomplished. A number of investigations have been made of the adjustment of rat, monkey, and chimpanzee to problems of this sort, and the characteristic outcome has been a marked initial negative transfer subsiding gradually in successive reversals.¹ Comparable experiments have not previously been made with submammalian forms, although there has been some interest in the repeated reversal of spatial habits by simpler animals.² The primary aim of the present work with the fish was to examine the possibility that the discrimination-reversal function might provide some basis for the analysis of phylogenetic differences in learning ability.

A secondary purpose of the present work was to explore the functional properties of a new technique for the study of discriminative learning in the fish. Various methods have been developed in the past, requiring the animal to enter one of

* Received for publication November 1, 1956.

¹ M. F. Fritz, Long-time training of white rats on antagonistic visual habits, *J. compar. Psychol.*, 11, 1931, 171-184; I. Krechevsky, Antagonistic visual discrimination habits in the white rat, *ibid.*, 14, 1932, 263-277; F. P. Gatling, The effect of repeated stimulus-reversals on learning in the rat, *J. compar. physiol. Psychol.*, 45, 1952, 347-351; H. W. Nissen, A. H. Riesen, and V. Nowlis, Delayed response and discrimination learning by chimpanzees, *J. compar. Psychol.*, 26, 1938, 361-395; J. T. Cowles, Food-tokens as incentives for learning by chimpanzees, *Compar. Psychol. Monogr.*, 14, 1937 (No. 71), 1-96; Riesen, Delayed reward in discrimination learning by chimpanzees, *ibid.*, 15, 1940 (No. 77), 1-54; more complex designs, involving not merely the same pair of stimuli but generalized reversal-sets, have been used in work with the monkey, e.g. H. F. Harlow, Studies in discrimination learning by monkeys: I. The learning of discrimination series and the reversal of discrimination series, *J. gen. Psychol.*, 30, 1944, 3-12.

² See an early study of the crawfish by R. M. Yerkes and G. E. Huggins (Habit formation in the crawfish, *Cambarus affinis*, *Harvard psychol. Studies*, 1, 1903, 565-577) and a more recent comparative investigation by Emanuel Seidman (Relative ability of the newt and the terrapin to reverse a direction habit, *J. compar. physiol. Psychol.*, 42, 1949, 320-327).

several compartments or to approach one of several baited stimulus-objects, but none has been flexible enough to come into wide use.³ For the present experiment, the lever-depression device of Haralson and Bitterman was adapted to the requirements of a two-choice discriminative problem.⁴ As Fig. 1 illustrates, a metal housing displaying two painted targets was lowered into S's living tank at the start of each trial. S's response, established by pre-training, was to strike at one or the other of the targets. A correct choice was rewarded with a small pellet of food dropped into the water, while an incorrect choice was unrewarded, and the

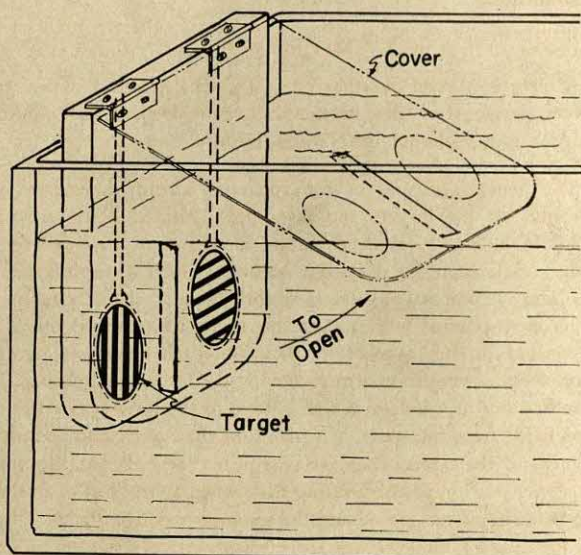


FIG. 1. AN APPARATUS FOR THE STUDY OF DISCRIMINATIVE
LEARNING IN THE FISH

Only certain general features of the apparatus and the position it occupies when lowered into S's tank are shown. The hinged cover of the housing is not, of course, lifted while the apparatus is in the water.

trial was terminated by the removal of the housing from the water. This apparatus has a number of advantages, of which the two most important may be specified. First, the criterion of response is completely objective—one of the targets must be displaced sufficiently to close an electrical circuit. Secondly, all handling of the fish, whether between trials or between sessions, is eliminated. In an attempt

³ See, for example, M. F. Washburn and I. M. Bentley, The establishment of an association involving color discrimination in the creek chub, *Semotilus atromaculatus*, *J. compar. neurol. Psychol.*, 16, 1906, 113-125; C. D. Reeves, Discrimination of light of different wave-lengths by fish, *Behav. Monogr.*, 4, 1919 (No. 9), 1-106; F. T. Perkins and R. H. Wheeler, Configurational learning in the goldfish, *Compar. psychol. Monogr.*, 7, 1930 (No. 31), 1-50; A. Meesters, Ueber die Organisation des Gesichtsfeldes der Fische, *Z. Tierpsychol.*, 4, 1940, 1-149.

⁴ J. V. Haralson and M. E. Bitterman, A lever-depression apparatus for the study of learning in fish, this JOURNAL, 63, 1950, 250-256.

to examine some of the characteristics of the new technique, about which little information was available beyond that acquired in a few preliminary runs, a methodological variable—correction versus noncorrection—was incorporated in the experiment. For one group of animals, each trial was terminated by the first response, whether it was correct or not; for a second group, opportunity for correction of errors was provided. The secondary purpose of the experiment was not, of course, unrelated to the first. If the methodological variable influenced rate of learning, a certain generality might be ascribed to any pattern of transfer common to the two training conditions.

METHOD

Subjects. Ten mature African mouthbreeders (6 females and 4 males, each about 2 in. long) were drawn from the stock-tanks of the Department of Animal Behavior of the American Museum of Natural History.

Apparatus. The general features of the apparatus already have been described. The metal housing for the targets and lever-arms, painted black, was about 5.5 in. wide, 7.5 in. high, and 1 in. deep. In its hinged front cover there were two circular openings, each 2 in. in diameter, between which a vertical fin projected about 0.75 in. The circular targets, each slightly greater than 2 in. in diameter, and in this experiment painted with 0.25-in. black and white stripes, were displayed through the openings in the cover of the housing. Each target was a light metal disk mounted on a light metal lever-arm pivoted at the top of the housing. The target slipped easily in and out of a holder at the bottom of the lever-arm. After each trial, the housing was lifted from the water, at which time the cover could be raised and the lateral arrangement of the targets changed as required. Sets of contact-points were so situated in the upper portion of the housing that, when *S* struck at or pushed against one of the targets, a circuit was closed and the response was signalled to *E* by the click of a relay. In the present work, *E* rewarded each correct response by dropping a small piece of preferred dry food into the water at the end of the tank opposite the housing; although provisions for automatic feeding might have been made without great difficulty,⁵ the effort did not appear to be warranted.

Preliminary training. Each *S* was housed in an individual tank, 10 in. long, 6 in. wide, and 8 in. high. All sides of the tank were opaque except one of the smaller ones, through which observations were made. After each *S* had adjusted to this mode of existence, taking food readily from the fingers of *E* during the daily feeding period, preliminary training was begun. Fish and tank were carried from the living area to the experimental setting and placed in a fixed position beneath the target-housing. After a few minutes, the housing was lowered into the tank. For purposes of pretraining, one of the openings in the cover was closed by means of a black metal disk, while the other displayed a gray target to which a bit of window-screening was fixed. (The gray target was shifted from side to side during the pretraining to prevent the development of spatial preferences.) Between the bit of screening and the gray target was a fragment of fresh beef liver which attracted the fish. When, in the course of nibbling at the liver, *S* dis-

⁵ W. A. Deterline, An automatic device for the investigation of operant behavior in the fish, this JOURNAL, 69, 1956, 291-295.

placed the target sufficiently to signal a response, *E* dropped a pellet of food into the water which *S* would then turn about to take. After several such trials, the liver was removed and the response was elicited by the unbaited target. When the response to the unbaited gray target was well established, experimental training was begun. The response transferred readily to the discriminanda.

This method of training has several advantages over the method of successive approximation employed by Haralson and Bitterman.⁶ It probably involves less variability from *S* to *S* and from *E* to *E*. Certainly it makes far less demand upon the skill of *E*. With provisions for automatic feeding, *E* need not even be present.

The food used in the experiment was a highly attractive dry mixture specially prepared at the American Museum of Natural History. It was in the form of irregularly shaped pellets, screened repeatedly for uniformity of size. Each pellet was about 1/32 in. in diameter.

Beginning with their isolation in individual tanks, the *Ss* were maintained on a ration of 20 such pellets per day, which was about 1/3 the number necessary to satiate the average fish. (The health of these animals seems to be impaired if they are given all the food they will eat.) Any part of the daily ration which was not earned during a given training session was given soon after *S* was returned to the living area.

Discriminative training. There were 20 massed trials per day, with the horizontally and vertically striped targets interchanged in position from trial to trial in accordance with orders selected from Gellermann.⁷ Half the *Ss* began with horizontal stripes positive, and half began with vertical stripes positive. Each trial started with the lowering of the housing into the tank and terminated with its withdrawal. For the 5 *Ss* assigned to the Noncorrection-Group, each trial was terminated by the first response, whether correct or not. For the 5 *Ss* of the Correction-Group, each trial was terminated either by the first correct response or by the fifth incorrect response—that is, if the first response was incorrect the housing remained in the water and *S* had opportunity to make as many as four additional incorrect responses before it was withdrawn. If a correct response was made before the fifth incorrect response, *S* was rewarded and the trial was terminated.

To forestall the possibility of extinction early in the reversal-training, a guidance-procedure was employed after any five consecutive unreinforced trials. Guidance consisted in the introduction of the positive target (the other opening in the housing being concealed, as in pretraining, by a black disk) and reinforcing the response which *S* would, sooner or later, make to it. The position of the positive target during guidance always was the same as it had been on the previous trial. The guidance-procedure was used with both groups, although only the Correction-Group showed any tendency toward extinction.

The criterion of learning was 17 errorless trials in each day's 20—that is, 17 trials on which the initial response was correct. (The *Ss* of the Noncorrection-Group had, of course, no opportunity for more than one response on each trial.) The day after each *S* reached criterion on the first problem (Reversal 0), training on the first reversal (Reversal 1) was begun, and continued to the same criterion. The experiment continued in this fashion for several months, during

⁶ Haralson and Bitterman, *op. cit.*, 250-256.

⁷ L. W. Gellermann, Chance orders for alternating stimuli in visual discrimination-experiments, *J. genet. Psychol.*, 42, 1933, 356-360.

which time all of the Ss achieved at least 8 reversals, and 7 of the 10 Ss achieved at least 13 reversals.

RESULTS

In Fig. 2, the performance of Correction- and Noncorrectoin-Groups is plotted in terms of mean initial errors. The curves are based on the data for the 7 Ss (4 in the Correction-Group and 3 in the Noncorrection-Group) that completed at least 13 reversals. (Very much the same pattern appears when the data for all Ss on 8 reversals are plotted.)

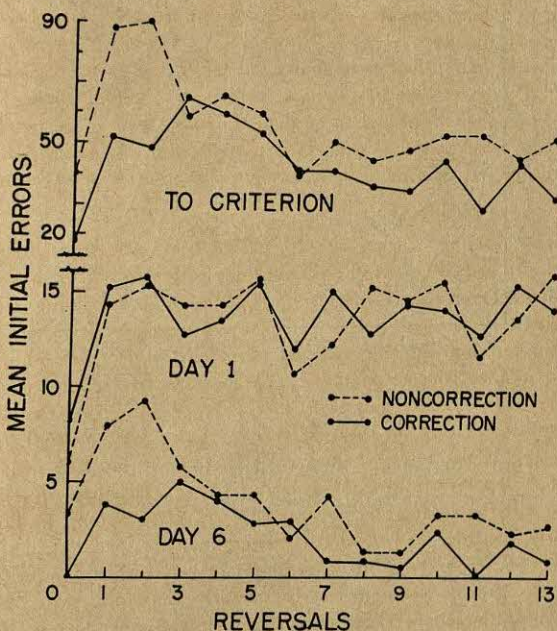


FIG. 2. MEAN INITIAL ERRORS PER REVERSAL

The top curves, showing mean initial errors to criterion for each reversal, are typical of those obtained with higher animals. They rise in the early reversals and then fall gradually. That is, there is high negative transfer at first, which then gradually grows less. Little improvement is evident here after the sixth reversal. It should be noted that the initial problem (Reversal 0) is more difficult for the noncorrected Ss—the best of the 5 noncorrected Ss made more errors than the poorest member of the Correction-Group. Interference develops more rapidly in the Noncorrection-Group, and the amount of interference is greater at first, although the difference in difficulty soon disappears.

The middle curves in Fig. 2 show mean initial errors for the first day—that is, for the first 20 trials—of each reversal, and the bottom curves show mean errors for the sixth day of each reversal. The curves of Day 6 clearly resemble the curves for errors to criterion, while those of Day 1 do not. The curves of Day 1 show no difference between the Correction- and Non-correction-conditions, and no change in the amount of interference from the first to the thirteenth reversal. Plots for Days 2, 3, 4, and 5 constitute a

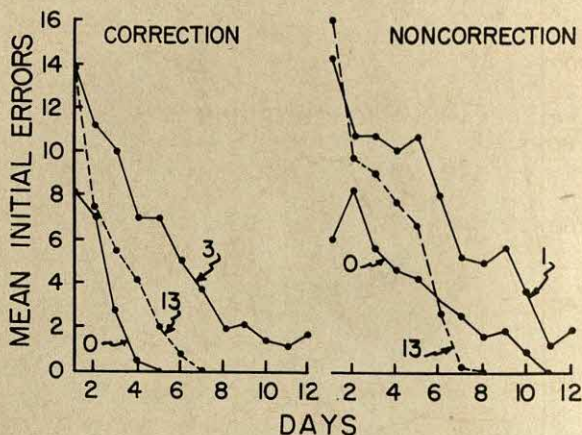


FIG. 3. LEARNING CURVES FOR SELECTED PROBLEMS

set of curves of intermediate shape, gradually becoming less similar to those for Day 1 and more similar to those for Day 6.

It is instructive to compare these data with those obtained by Gatling for the rat.⁸ Gatling's curve for errors to criterion has the same general shape as the criterion-curves for the fish—it rises early and then falls in negatively accelerated fashion. The very same shape is found, however, in Gatling's curve of mean errors on the first 10 trials of each reversal, as well as his curve for mean errors on trials 11-20 of each reversal, and for mean errors on Trials 41-50 of each reversal. In the rat, it seems, *the general pattern of recovery from interference found in successive reversals is reflected at all stages of each reversal*. In the fish, by contrast, *the over-all pattern of recovery must be traced to effects which appear only in the later stages of each reversal*.

Another way of looking at the pattern of interference is to plot learning curves for selected reversal-problems. The set of curves in the left-hand part of Fig. 3 is for the original problem (Reversal 0), the last reversal (Reversal 13), and the third reversal (the most difficult) of the Correction-Group. The increase in difficulty from Reversal 0 to Reversal 3 is reflected at all stages of training, while the improvement from Reversal 3 to Reversal 13 is reflected only after the first day and

⁸ Gatling, *op. cit.*, 347-351.

increasingly thereafter. The set of curves in the right-hand portion of Fig. 3 is for Reversal 0, Reversal 13, and Reversal 1 of the Noncorrection-Group. (Although the second reversal was the problem of maximal difficulty for this group, the curve of the first reversal is shown for formal reasons. Since the last reversal was odd-numbered, progress is indicated by comparison with an odd-numbered early reversal—that is, an early reversal with the same positive stimulus for each *S*. Actually, the curves for Reversals 1 and 2 are almost identical.) The same pat-

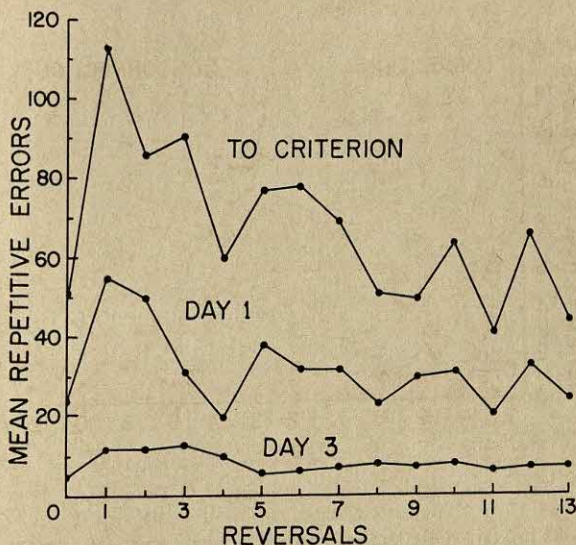


FIG. 4. MEAN REPETITIVE ERRORS PER REVERSAL IN THE CORRECTION GROUP

tern of change is shown by the curves of the Noncorrection-Group as by those of the Correction-Group. From Reversal 0 to Reversal 1, there is an increase in difficulty at all stages of learning, while the recovery from negative transfer (between Reversal 1 and Reversal 13) appears as a change in rate of error-elimination from a fixed first-day point. Plotted in this way, Gatling's data for the rat would yield a series of parallel curves; the curves for early and late reversals would differ in over-all level rather than in rate of fall.

Another pattern of improvement in the fish is revealed in the elimination of repetitive errors by the Correction-Group. The top curve of Fig. 4 shows mean repetitive errors to criterion for each reversal; the middle curve shows mean repetitive errors on the first day of each reversal; and the bottom curve shows mean repetitive errors on the third day of each reversal. (It will be remembered that an *S* could make as many as four repetitive errors on each trial.) These curves indicate that the principal change from reversal to reversal occurs on the first day of each problem. Unfortunately, no comparable data are available for other forms. The primary significance of these curves for the fish at the present time lies in the fact that they afford a further contrast with the kind of change found for initial errors.

From the comparative point of view which led to this experiment, the central finding relates to the different patterns of initial-error elimination found in fish and rat. In the data for the fish, recovery from negative transfer can be represented by a family of learning curves which differ in slope. In the data of Gatling for the rat, recovery can be represented by a family of parallel-curves. It is much too early, of course, to conclude that a species-difference in learning has been isolated in these studies. It may be, for example, that the results are due only to a difference in some peripheral variable, such as drive—that a change in the drive-level of the fish would have produced a transfer-function like Gatling's, or that a change in the drive-level of the rat would have produced a transfer-function like that found here in the fish. What, too, of the difficulty of the original discrimination? Would the use of a different pair of stimuli have produced a more rat-like function in the fish? The only way to answer such questions is to extend the work on both forms with wide variation in the values of important parameters. It is interesting to note in this connection that the two training methods here employed—correction and noncorrection—though yielding significant differences in rate of learning, give the same unique pattern of transfer.

If the results thus far obtained hold up under a broad range of conditions, we can perhaps begin to answer a question on which there has been a good deal of speculation but little research in the past. It was Hull's guess, for example, that equations of the same form would be found to describe the behavior of all species, although the equations might differ with respect to the value of constants—as the value of g in $s = \frac{1}{2} g t^2$ varies from point to point on the surface of the earth.⁹ The difference uncovered in the present experiment seems, however, to be qualitative rather than quantitative. If substantiated by future research, it would suggest that when we go from fish to rat we need, not new constants, but new equations.

SUMMARY

The course of reversal-learning in the fish was studied under conditions analogous to those employed in work with higher forms. Analysis of initial errors showed, both in a Correction- and a Noncorrection-Group, a pattern of recovery from negative transfer qualitatively distinct from that previously obtained for the rat, a finding whose significance for the problem of phylogenetic differences in learning ability remains to be assessed. Analysis of repetitive errors in the Correction-Group yielded still another pattern of recovery.

⁹ C. L. Hull, The place of innate individual and species-differences in a natural-science theory of behavior, *Psychol. Rev.*, 52, 1945, 55-60.

VISUAL BEATS CAUSED BY SIMULTANEOUS ELECTRICAL AND PHOTIC STIMULATION

By JOHS. CLAUSEN and CHRISTA VANDERBILT,
The Training School at Vineland

In acoustics, beats occur when the ear is stimulated by two tones of slightly different frequencies. The theory for the phenomenon, which is generally accepted, was originated by Helmholtz. It states that auditory beats result from imperfect analysis of the cochlear mechanism—that the components, the frequencies of the stimuli, affect overlapping sections of the basilar membrane.¹ A similar phenomenon seems to exist for vision. When the eye is stimulated by an alternating electrical current through electrodes placed on the surface of the skin, this stimulation results in a sensation of flickering light (phosphenes).² If an observer's eye is stimulated by an alternating current while looking at an intermittent photic light, the two stimuli interact. The result of this interaction is a visual phenomenon which has the essential characteristics of beats as we know them from acoustics.

The following is a typical set of conditions that yield the visual beat phenomenon. Electrical source: 20 ~ of 200-300 μ A. applied to a pair of electrodes attached to the forehead and the temple. Photic source: Strobotac set at 21 ~ and as close as possible to the eye that the flicker will occupy the entire visual field. The resulting experience, superimposed on the flicker from the electrical and the photic stimulation, is a light or pattern of light which moves rhythmically in the horizontal plane across the field from one extreme of the field to the other and back again. The frequency of the movement is the difference between the frequency of the electrical stimulation (20 ~) and the photic stimulation (21 ~), i.e. 1 ~.

To the best of our knowledge only one study, that by Schwarz, mentions the phenomenon of visual beat.³ Schwarz reports in passing that simultaneous electrical and photic stimulation of the eye produces beats in a frequency-range of 8 to 20 ~ when the frequency-difference between the two stimulus sources is small. In the experiment on phosphenes, mentioned above, Clausen observed beats with simultaneous electrical and photic stimulation of the eye.⁴ On this background we proposed to deter-

* Received for publication October 11, 1956. This investigation was supported by Research Grant NSF-G2106 from the National Science Foundation.

¹ J. C. R. Licklider, Basic correlates of the auditory stimulus, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 1015.

² Johs. Clausen, *Visual Sensations (Phosphenes) Produced by AC Sine Wave Stimulation*, 1955, 32 ff.

³ Friedrich Schwarz, Ueber die Wirkung von Wechselstrom auf das Sehorgan, *Z. Sinnesphysiol.*, 67, 1938, 227-244.

⁴ Clausen, *op. cit.*, 32 ff.

mine the conditions under which visual beats would occur, particularly with regard (1) to frequency-range of the two stimulus-sources; (2) to frequency-difference between them; and (3) to the intensity of electrical and photic stimulation. If this visual phenomenon can properly be termed beats, it is a basic requirement that its frequency should equal the difference between the frequencies of the electrical and photic sources.

APPARATUS AND PROCEDURE

The source of electrical stimulation was a Hewlett-Packard Low Frequency Oscillator (Model 202C). Silver electrodes, 25 mm. in diameter, were attached to the forehead and the right temple of the observer.⁵ In those cases where current-intensity was an experimental variable, the current was recorded by a Hewlett-Packard Low Frequency Vacuum Tube Voltmeter (Model 400B) with a shunt resistor. In the majority of the experiments a General Radio Strobotac (Model 631 BL) was the source of intermittent light. When frequency stability was of particular concern, we used a Grass Photo Stimulator (Model PS1) as the light source. When the Strobotac was used, various luminance levels of the light-stimulus were attained by inserting Eastman Kodak Wratten Neutral Density Filters in front of the light source, and lights of different wave lengths were obtained by inserting Eastman Kodak Narrow-Band Filters in front of the source. *O* sat in front of the light-source and looked at the light with the right eye, across which the electrodes had been placed. The left eye was covered with an eye patch. As mentioned in the introduction, the main characteristic of visual beats is a rhythmic movement back and forth across the field. To determine the beat-frequency *O* counted the number of complete cycles of the movement, *i.e.* from the right to the left and back again for a period of 30 sec.

With regard to statistical analysis of our frequency-counts we are faced with a rather unusual situation. Our hypothesis about beat-frequency is that it equals the difference between the frequencies of the electrical and the photic stimuli. This can be demonstrated only if both stimulus-sources have a high degree of stability. A Hewlett-Packard Electronic Counter (Model 522B) revealed that the sine wave oscillator maintained its frequency in a satisfactory fashion, but that both the Strobotac and the Photo Stimulator varied enough in frequency to influence the beat-counts. The instruments employed drifted somewhat in an unpredictable fashion. When the stability was high, as indicated by the Electronic Counter, the beat-counts were in exact accordance with the hypothesis. When the Electronic Counter indicated definite frequency-changes, the beat-frequency could be off as much as 3 or 4 counts. If one under such circumstances should accumulate data and compute means and standard deviations, these measures would reflect on the stability of the instruments rather than the visual phenomenon we are trying to study. The procedure we chose was to follow continuously the period-recording of the photic stimulator on the Electronic Counter during the experiment. This gave us a chance to adjust the photic stimulator before each count, even if we could not adjust it during a count. We accepted a fluctuation of the instruments which caused a devia-

⁵ For further details about electrode attachment and procedure see Clausen, *op. cit.*, 27-31.

tion of no more than 3.3% from the expected or predominant count. Occasional counts which, due to such instability were off more than this, were discarded.

RESULTS

(1) *Beat-frequency: Experiment I.* The Grass Photo Stimulator was used to produce frequencies, ranging from 5 to 50 ~ in steps of 5 ~. The electric source was adjusted to a frequency of 1 ~ below and to 1 ~ above the frequency of the photo stimulator. The intensity of the electric stimulus was so adjusted that the phosphenes appeared reasonably clear at all frequencies (300-500 μ A.). Each *O* made five counts for each of the photic stimulus-frequencies with electric frequency 1 ~ below and five counts with electric frequency 1 ~ above the photic frequency. Beats were observed in the frequency-range from 5 to 45 ~. Within the 5-45 ~ range each *O* made a total of 90 counts which were within the established criteria of 29 to 31 counts per each 30-sec. counting-period. Below 5 ~ the beats were not observed since the phosphene threshold in this range is too high to be attained without appreciable discomfort. Above 45 ~ the photic stimulus appears fused, and beats did not occur. It seems that both stimulus-sources must actually produce flicker in order for beats to occur.

Experiment II. Having established the fact that the frequency of the observed beats corresponds to the difference between the frequencies of the two stimulus-sources over the entire range where visual beats could be produced, we proceeded to vary the frequency-difference between the two sources. In this experiment the electrical stimulus was kept at 20 ~ and the frequency of the Photic Stimulator varied from 14 to 26 ~ in steps of 0.5 ~. In the frequency-range from 17.5 to 22.5 ~ it is possible to count the beats. Five counts were made by each *O* at each of the 10 frequency-combinations making a total of 50 counts. Within the 3.3% limits the counts were equal to the difference between the two component frequencies. In the range of 2.5 to 4 ~ frequency-difference, beats could be observed but were too fast to be counted accurately. With even greater differences one of the *Os* reported relatively slow beats which may be related to multiples or submultiples of the fundamental frequency.

Experiment III. That multiples and submultiples of the fundamental frequency actually produce beats was demonstrated in an experiment where the following frequency-combinations were used: photic stimulations of 15, 10, 40, and 30 ~; with electrical stimulations of 29 and 31, 19 and 21, 19 and 21, and 14 and 16 ~, respectively.

In each of the eight stimulus-combinations, five counts were made by each of the two *Os*. Within our stated criterion all counts showed beat-frequency of 1 ~.

In summary, then, it appears that in the frequency-range from 5 to 45 $\sim$ and at a frequency-difference up to 4 $\sim$, visual beats occur at frequencies equalling the difference between the electrical and the photic stimulations. Multiples and submultiples of the fundamental frequency also produce beats.

(2) *Stimulus-intensity: Experiment I.* The purpose of this experiment was to establish the relationship between stimulus-intensity and beat-frequency. Frequencies of 10, 15, 20, 25 and 30 $\sim$ of the electrical stimulus were selected, with the frequency of the corresponding photic stimulation 1 $\sim$ higher. In half of the experiments we kept the intensity of the photic stimulus constant at an intermediate level (1.5 log filter), and varied the intensity of the electrical stimulus. For each frequency we started with a current intensity of 700 μ A., decreasing in steps of 50 μ A. until beats could no longer be seen, making one count for each intensity-level of the current. This procedure was repeated three times for each *O*. A total of 190 counts were made by the two *O*s. Within the accepted limit the beat-frequency equaled the difference between the frequencies of the two stimuli, indicating that beat-frequency is independent of current intensity.

In the other half of the experiment we kept the electrical stimulus constant at about 350 μ A. and varied the luminosity of the photic stimulus. This was done by inserting neutral density filters in front of the light-source in an increasing series in steps of 0.5 log units, until the beats could no longer be observed. Again we made one count at each filter-level, and again we repeated the procedure three times for each *O*. A total of 167 counts were made. As in the first part of the experiment the counts were within the accepted limits, indicating that beat-frequency is also independent of luminosity.

Experiment II. This experiment was designed to determine the minimal intensity of the electrical stimulus necessary to produce beats at frequencies from 5 to 45 $\sim$ in steps of 5 $\sim$. A frequency difference of 1 $\sim$ between the electrical and the photic stimulus was used. A log 1.0 filter was placed in front of the Strobotac and this level of luminosity was maintained throughout the experiment. At each frequency we started with a current-intensity of 700 μ A. and decreased in steps of 50 μ A. As the beats were too feeble to be counted at the lowest current-intensities, we discarded the counting and had the *O* report whether or not he saw the beats. Each *O* repeated the procedure six times, with systematic rotation of the series-position of the frequencies in order to avoid the effect of

position. The combined results from the two *O*s are presented in Fig. 1 where each point represents an average of 12 determinations. This curve, with its U-shaped form and lowest value at 20 ~, has the same general form as is obtained when phosphene threshold is plotted against frequency.⁶ The main difference is that the curve from the present experiment is shifted upwards on the intensity-scale. As the phosphenes at threshold value appear in the peripheral part of the field, and only at higher stimulus-intensities occupies an appreciable part of the field, it is indicated that the phosphenes must appear centrally in the visual field in order to yield this particular interaction with the photic stimulus.

Experiment III. In an experiment similar to the preceding one, we tried

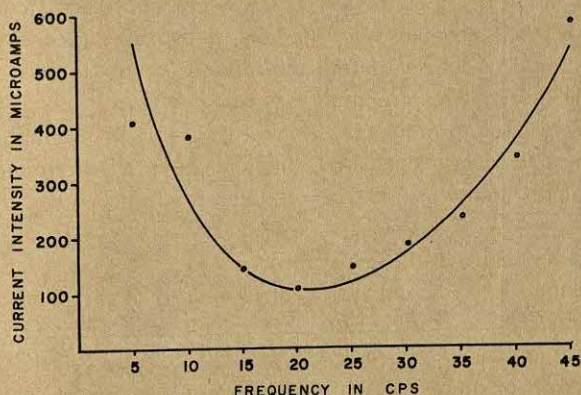


FIG. 1. RELATIONSHIP BETWEEN MINIMAL INTENSITY OF ELECTRICAL STIMULATION REQUIRED TO PRODUCE BEATS AND FREQUENCY OF THE STIMULI

to determine the minimal luminance-level of the photic stimulus necessary to produce beats. The frequency-range was from 5 to 50 ~ in steps of 5 ~. The intensity of the electrical stimulus was kept constant at 350 μ A. The luminance-level was varied in a decreasing fashion by inserting neutral density filters in steps of 0.5 log units in front of the light-source until the beats disappeared. The procedure was repeated five times for each of the two *O*s, with systematic rotation of the series position of the frequencies. The results are presented in Fig. 2 where each point is an average of 10 determinations. This curve has a striking similarity to intensity-frequency plots of critical flicker fusion, including a break

⁶ Clausen, *op. cit.*, 58.

around 25 $\sim$, customarily referred to as the rod-cone break. Fig. 2 shows that at a low frequency of the intermittent light only a faint light (high filter-value) is necessary to produce the beats. At high frequencies a high luminance-level is necessary. This is consistent with findings from CFF determination; namely, that bright light has a higher flicker threshold than has dim light. Thus it appears that the intermittent light has to be perceived as flickering in order to produce beats.

Experiment IV. The purpose of this experiment was to determine whether there is any interaction between intensities of electrical and

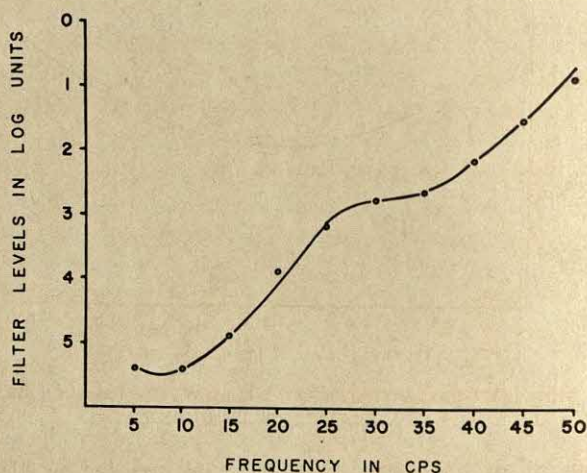


FIG. 2. MINIMAL LUMINOSITY-LEVEL OF PHOTIC STIMULUS AS RELATED TO FREQUENCY IN PRODUCING BEATS

photic stimulation. Frequencies employed were 24 and 25 $\sim$ for the electrical and photic stimuli, respectively. The photic stimulus was varied by inserting filters from 0.0 to 2.5 log units in steps of 0.5 log units. For each of these filter-levels we varied the current-intensity, starting at 250 μ A. and reducing it in steps of 50 μ A. until the beats disappeared. With seven repetitions for the two Os, a total of 84 determinations were made. Critical current-intensity plotted against filter-levels results in a slightly U-shaped curve presented in Fig. 3, where each point is an average of 14 determinations. The curve indicates an interaction between the intensity of the two stimuli, in that beats occur with minimal current-intensity at an intermediate luminosity of 1.5 log unit filter.

Summarizing the essential findings about stimulus-intensity it may be

stated that the electrical stimulus has to be well above phosphene threshold, and for any frequency of the intermittent light, the intensity has to be sufficient to make the resulting sensation appear as flickering if beats are to occur. An interaction between the two stimulus-intensities is also indicated.

As supplements to those described above, we made a few additional experiments. As they were pilot studies rather than systematic experiments, the results will only be mentioned in summary fashion. (1) By inserting Eastman Kodak Color filters

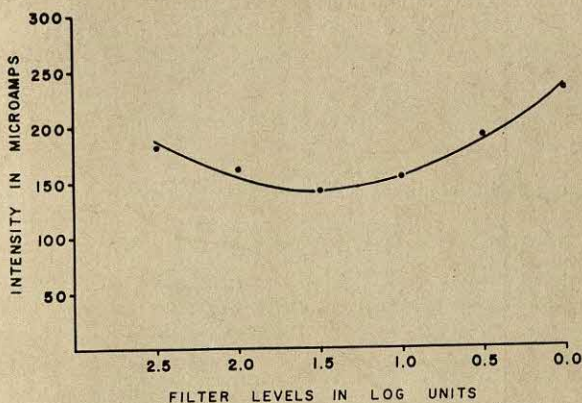


FIG. 3. MINIMAL CURRENT-INTENSITY NECESSARY TO PRODUCE BEATS AT VARIOUS LEVELS OF LUMINOSITY

Nos. 70, 72B, 73, 74, 75 and 76, it was found that they had no other effect on beats than that related to their transmission factor of luminosity. (2) Beats were observed when the eye was stimulated by two electrical sources; namely, a sine wave and a square wave oscillator. In this situation beats were also experienced from the accompanying cutaneous sensation. This sensation waxed and waned in intensity corresponding to the difference in frequency between the two stimuli. (3) When the eye was stimulated by intermittent flashes at a rate of 20 per sec. and the ear was stimulated by audible clicks of 19 per sec., the two resulting sensations remained separate without any indication of rhythmic phenomena resembling beats.

DISCUSSION

In a previous study concerning electrical phosphenes, it was concluded that the weight of the evidence points to neural structures of the retina (the neural part of the visual receptors, the bipolar cells, or both) as being the locus of stimulation for the electrical current.⁷ If this is the case we may assume that the visual beats occur as an interaction between the

⁷ Clausen, *op. cit.*, 81.

electrical stimulation and the electrical activity of these structures resulting from the photic stimulation. We may further assume that this interaction involves the rhythmical bursts of spike potentials rather than the individual spikes. The bursts would have a rhythm corresponding to the frequency of the intermittent light, as indicated by Granit's recordings from single nerve fibers.⁸ Consistent with this theory is the fact that intensity of the light-stimulus does not influence the beat-frequency, whereas it determines the frequency of the action potentials.

Occurrence of rhythmical bursts is probably more characteristic for artificial stimulation in the laboratory than it is for natural stimulation. Thus the beat-phenomenon we have observed may not be a common type of nervous function. It might, however, be hypothesized that periodical discharges in two neurons will produce beats within a certain range of frequency-differences.

As indicated by the results, visual beats occur only when the intermittent light is perceived as flicker. Only under these circumstances will the bursts be sufficiently distinct to interact with the direct electrical stimulation. This limitation might be accounted for by the clash of on and off stimuli at relatively high frequencies of photic stimulation as demonstrated by Granit.⁹

Wever asserts that auditory beats pass through four quality steps as the difference between the frequencies of the two components is increased. In our experience there are no such quality changes for visual beats.¹⁰ Also the auditory beats appear to exist for greater frequency-differences between the components than do visual beats. The indication, then, is that the two types of beats have different underlying mechanisms. Helmholtz assumed that auditory beats involve overlapping segments of the basilar membrane,¹¹ whereas the visual apparatus does not have a similar physical structure. Thus we may conclude that our findings indicate that beats can be produced by interaction of stimuli in neurons.

SUMMARY

In a series of experiments we have explored the stimulus-characteristics which will produce visual beats by simultaneous photic and electrical stimulation of the eye.

⁸ Ragnar Granit, *Receptors and Sensory Perception*, 1955, 169.

⁹ Granit, *op. cit.*, 169.

¹⁰ E. G. Wever, Beats and related phenomena resulting from the simultaneous sounding of two tones, *Psychol. Rev.*, 36, 1929, 402-418, 512-523.

¹¹ Licklider, *op. cit.*, 1015.

The beat-frequency was found to equal the difference between the frequencies of the two component stimuli, and to be independent of stimulus-intensity. Beats were produced in a frequency-range of 5 to 45 ~ and with a maximal difference of 4 ~ between the frequencies of the two stimuli. When the minimal intensity of electrical stimulation necessary to produce beats is plotted against frequency, the curve is similar to the phosphene threshold-curve, and a corresponding plot for the photic stimulation results in a curve similar to that of *CFF*. Stimulation by two electrical sources results in beats, whereas stimulation by intermittent light and audible clicks does not. It is concluded that visual beats are produced by interaction of stimuli in neurons.

WORD-ASSOCIATION AND SEMANTIC DIFFERENTIATION

By SAMUEL J. MESSICK, Educational Testing Service, and
CHARLES M. SOLLEY, The Menninger Foundation

In a previous article the authors related one measurement of meaning to certain hypotheses about the formation of meaning.¹ A statistical model for concept-formation was outlined in which the meaning of a concept was considered to be a reflection of probability-relationships among its referents. Any potential combination of the perceptual events related to a concept was viewed as a possible referent. The likelihood that a concept would refer to a particular referent depended upon the probability of occurrence of specific combinations of events. Concept-formation was viewed as a process whereby an individual, on the basis of more and more experience with more and more combinations of related events, abstracted communalities from the population of referents.

Probability-learning procedures were used to study the acquisition of perceptual characteristics for an artificial concept for which the 'true' distribution of referents was known.² The concept used was 'Tribe,' and the referents for this concept were 20 drawings of little men whose perceptual characteristics varied as to whether they were *tall* or *short*, *happy* or *sad*, *fat* or *skinny*, and *black* or *white*. Distributions of perceptual characteristics for two tribes, Tribes A and B, are given in Table I. This table indicates that the two tribes are identical in their marginal frequencies for the different stimulus-characteristics; *i.e.* there are 15 tall and 5 short stickmen in each tribe, 12 happy and 8 sad, 16 skinny and 4 fat, and 20 white. The two tribes differ, however, with respect to the joint occurrence of combinations of these perceptual characteristics; *e.g.* in Tribe A there are 12 tall, skinny, happy, white stickmen, while only 3 have that description in Tribe B. Thus, with respect to the marginal frequencies for each perceptual characteristic, the two tribes have identical 'meanings', but with respect to the joint occurrence of combinations of these characteristics, the tribes have somewhat different 'meanings.'

In the previous experiment probability-learning procedures were used with these two artificial tribe-concepts. It was found that Ss learned the joint probabilities for

* Received for publication March 13, 1956. This research was completed while the authors were at the Menninger Foundation as USPHS Postdoctoral Research Fellows of the National Institute of Mental Health.

¹ C. M. Solley and S. J. Messick, Probability-learning, the statistical structure of concepts, and the measurement of meaning, this JOURNAL, 70, 1957, 161-173.

² W. K. Estes and C. J. Burke, A theory of stimulus-variability in learning, *Psychol. Rev.*, 60, 1953, 276-286; H. W. Hake and R. Hyman, Perception of the statistical structure of a random series of binary symbols, *J. exp. Psychol.*, 45, 1953, 64-74.

the occurrence of combinations of characteristics and not just the marginal aspects of the concepts alone. Since the marginal probabilities may be generated from the joint probabilities, the Ss had acquired information about two statistical aspects of the meaning of these concepts—the marginal probabilities for individual characteristics and the joint probabilities for combinations of characteristics.

It was of interest to see which of these statistical aspects would be reflected by instruments designed to measure meaning. If an instrument were to reflect only marginal aspects of a concept, then Tribes A and B would be identical with respect to measured meaning. If it were to reflect joint occurrences of characteristics or some combination of joint occurrence plus marginal characteristics, then the meanings of Tribes A and B would be differentiated. In using the semantic differential as a measure of meaning, Tribes A and B were found to be identical, and it was concluded that this instrument was reflecting only the marginal characteristics of a concept's meaning.³ It was believed, however, that a modification in the semantic differential procedures might lead to a differentiation in the meanings of A and B.

TABLE I
FREQUENCY OF OCCURRENCE FOR VARIOUS COMBINATIONS OF STIMULUS-CHARACTERISTICS IN TWO HYPOTHETICAL TRIBES

Combinations	Tribe		Combinations	Tribe	
	A	B		A	B
tall-skinny-happy-white	12	3	tall-skinny-happy-black	0	0
tall-fat-happy-white	0	4	tall-fat-happy-black	0	0
tall-skinny-sad-white	3	8	tall-skinny-sad-black	0	0
tall-fat-sad-white	0	0	tall-fat-sad-black	0	0
short-skinny-happy-white	0	5	short-skinny-happy-black	0	0
short-fat-happy-white	0	0	short-fat-happy-black	0	0
short-skinny-sad-white	1	0	short-skinny-sad-black	0	0
short-fat-sad-white	4	0	short-fat-sad-black	0	0

If combinations of characteristics were included in the task directly as adjectives modifying the concept to be rated, the semantic differential might also reflect the joint probabilities. That is, instead of rating 'stickmen' on *tall-short*, *happy-sad*, etc., Ss could be asked to rate 'tall stickmen' on *happy-sad*, *fat-skinny*, etc.

The question now arises as to whether there are other measures of meaning which adequately reflect joint probabilities as well as marginal characteristics. In the present experiment a word-association technique was used as a possible index of meaning to see which statistical aspects of concepts it would reflect.

Word-association. It has long been observed in word-association tests that if a stimulus-word A is given, S might respond with B, but if B is given, the response is not necessarily A. In other words A might imply B, but B does not necessarily imply A. This observation might be conceptualized in terms of a geometric model as a situation in which the psychological distance, or perhaps the semantic distance,

³ C. E. Osgood, The nature and measurement of meaning, *Psychol. Bull.*, 49, 1952, 192-237; Solley and Messick, *op. cit.*, 164-172.

from A to B is not necessarily the same as the psychological distance from B to A. Such a difference would not be reflected in the Euclidean model underlying the semantic differential.⁴ Word-association tests, however, might very well reflect such directional aspects of meaning and thus provide a possible method for studying asymmetric meaning relationships among a concept's referents.

A possible mathematical model for the analysis of word-association data may be derived from a consideration of the conditional probabilities among a concept's referents. For example, in Tribe A consider the probability of occurrence for the characteristic *happy* given the occurrence of the characteristic *tall*. Since 12 of the 15 tall stickmen in Tribe A were happy, the conditional probability for the occurrence of *happy* given tall would be 0.80. The conditional probability for the occurrence of any one of the characteristics, X , given the occurrence of another one, Y , can be computed from $P(X|Y) = P(X,Y)/P(Y)$, where $P(X|Y)$ is the probability of X given Y , $P(X,Y)$ is the joint probability for the occurrence of X and Y , and $P(Y)$ is the marginal probability for the occurrence of Y providing $P(Y)$ is not zero.

Since the joint probabilities $P(X,Y)$ and the marginal probabilities $P(Y)$ are known for each of the characteristics in Tribes A and B, the conditional probabilities for the occurrence of one attribute given another can be computed. These theoretical or expected conditional probabilities for each pair of characteristics are presented in Tables II and III for Tribes A and B, respectively. The probabilities presented in these tables were so scaled that the sum of each row would equal unity, and the tables would thus represent stochastic matrices. Each of these matrices of expected conditional probabilities may be considered to be a transitional state in a Markov chain, although Markov analysis will not be attempted with the present data.

Since there is evidence that during probability-learning procedures Ss acquire both the joint probabilities and the marginal probabilities for each of the characteristics, it is theoretically possible for Ss' behavior to reflect the conditional probabilities described above.⁵ The following experiment was performed to see whether a word-association test involving the characteristics directly would adequately reflect these conditional probabilities. If they are reflected, the test must at the same time reflect the joint probabilities for the occurrence of combinations of characteristics, and thus the meaning of Tribes A and B would be differentiated. It is also possible, however, for word-association to differentiate Tribes A and B on some other basis and not follow the simple model of conditional probability described above. For such a case, however, it could also be concluded that the joint probabilities are reflected by the instrument. If the meanings of Tribes A and B are not differentiated, then either the marginal characteristics of the concept are being measured or else artifacts are being reflected.

Method: The Ss in this experiment were 18 women (secretarial students) from 18 to 21 yr. of age. Nine Ss were given Tribe A in a probability-learning procedure, and nine were given Tribe B.

The materials used in the learning situation were the two tribes described in

⁴ Messick. Some recent theoretical developments in multidimensional scaling, *Educ. psychol. Measmt.*, 16, 1956, 82-100.

⁵ Solley and Messick, *op. cit.*, 171-172.

Table I. Each tribe was composed of 20 cards in a random order, upon each of which appeared a schematized human figure with a certain combination of perceptual characteristics.

A probability-learning sequence was used to familiarize *S* with the perceptual characteristics of one of the tribes, and then a word-association test was given as a possible measure of the characteristics learned. In the probability-learning phase of the experiment each *S* was instructed to guess the characteristics of the top card when the deck was face down; e.g. he would guess whether the top drawing was *tall*, or *short*, then whether it was *happy* or *sad*, *fat* or *skinny*, and *black* or *white*. The card was then turned over, allowing *S* to scan the picture for approximately 5 sec. The card was then removed from view, and *S* made four guesses about the next card, one guess for each of the bipolar characteristics. When guesses had been made for all 20 pictures in this way, a different random order of the same deck was presented and the procedure repeated. Each *S* thus had an opportunity to scan the same deck twice. In the previous study in which the decks had been presented to *S* four times, it had been found that most of the learning had occurred by the end of the second presentation.

In the word-association procedure, *Ss* were instructed to think about the tribe of stickmen they had just seen. Then they were told that *E* would give a stimulus-word to which they should respond with the first words that came to mind. Each *S* was then presented with a series of 16 stimulus-words, composed of two random orders of the eight characteristics tall, short, happy, sad, fat, skinny, black, and white. All of *S*'s responses to these words were recorded, but it was agreed before the experiment to ignore responses that were opposite in meaning to the stimulus-word. Thus, if the stimulus was *white*, the response *black* was not included in the data, since it was felt that the response-tendency to give opposites in word-association tests might overwhelm other effects. Responses which were synonymous with, or descriptive of, one of the attributes were counted as responses for the attribute; e.g. for a given stimulus-word, responses of *happy*, *carefree*, *gay*, *smiling*, etc., were tallied as *happy*.

RESULTS

The responses to each stimulus-word were tallied and the proportions of these responses reflecting each perceptual characteristic were recorded. These obtained proportions for Tribes A and B are presented in Tables II and III, respectively; the proportions were again so scaled that the sum of each row would equal unity. These tables also give the expected marginal probabilities for each characteristic in the two tribes. The expected marginals are identical for Tribes A and B; they represent the proportions of times each characteristic appeared. Observed marginal probabilities were computed by finding the proportion of responses for one aspect of a bipolar attribute out of the total number of responses to that attribute elicited by all stimulus-words; e.g. the observed marginal for the characteristic *tall* was found by dividing the number of *tall* responses elicited by all eight

TABLE II
EXPECTED (E) AND OBTAINED (O) PROPORTIONS FOR WORD-ASSOCIATION RESPONSES
FOLLOWING PROBABILITY LEARNING OF TRIBE A

Stimulus- words		Responses							
		tall	short	fat	skinny	happy	sad	black	white
tall	E	.25	—	.00	.25	.20	.05	.00	.25
	O	.11	—	.04	.42	.27	.04	.00	.11
short	E	—	.25	.20	.05	.00	.25	.00	.25
	O	—	.24	.36	.04	.12	.20	.00	.04
fat	E	.00	.25	.25	—	.00	.25	.00	.25
	O	.04	.22	.26	—	.09	.35	.00	.04
skinny	E	.24	.01	—	.25	.19	.06	.00	.25
	O	.26	.07	—	.19	.33	.11	.00	.04
happy	E	.25	.00	.00	.25	.25	—	.00	.25
	O	.23	.09	.05	.14	.32	—	.00	.18
sad	E	.09	.16	.12	.12	—	.25	.00	.25
	O	.10	.25	.20	.10	—	.30	.00	.05
black	E	—	—	—	—	—	—	—	—
	O	.09	.00	.00	.09	.18	.27	.36	—
white	E	.19	.06	.05	.20	.15	.10	—	.25
	O	.12	.08	.08	.12	.32	.08	—	.20
Marginal probabilities*	E	.75	.25	.20	.80	.60	.40	.00	1.00
	O	.50	.50	.46	.54	.59	.41	.32	.68

* Marginal probabilities were not scaled to make the sum of the rows equal unity. They represent the proportion of times each characteristic appeared in a tribe.

TABLE III
EXPECTED (E) AND OBTAINED (O) PROPORTIONS FOR WORD-ASSOCIATIONS
FOLLOWING PROBABILITY LEARNING OF TRIBE B.

Stimulus- words		Responses							
		tall	short	fat	skinny	happy	sad	black	white
tall	E	.25	—	.07	.18	.12	.12	.00	.25
	O	.18	—	.00	.42	.21	.06	.03	.09
short	E	—	.25	.00	.25	.25	.00	.00	.25
	O	—	.26	.16	.19	.29	.06	.00	.03
fat	E	.25	.00	.25	—	.25	.00	.00	.25
	O	.36	.00	.27	—	.33	.00	.00	.03
skinny	E	.18	.07	—	.25	.12	.12	.00	.25
	O	.19	.24	—	.29	.09	.14	.00	.04
happy	E	.15	.10	.08	.17	.25	—	.00	.25
	O	.06	.18	.12	.03	.55	—	.00	.06
sad	E	.25	.00	.00	.25	—	.25	.00	.25
	O	.27	.05	.05	.18	—	.40	.00	.05
black	E	—	—	—	—	—	—	—	—
	O	.00	.14	.00	.14	.14	.00	.57	—
white	E	.19	.06	.05	.20	.15	.10	—	.25
	O	.17	.00	.00	.17	.17	.00	—	.50
Marginal probabilities	E	.75	.25	.20	.80	.60	.40	.00	1.00
	O	.60	.40	.14	.86	.76	.24	.31	.69

stimulus-words by the sum of *tall* plus *short* responses. An examination of these marginals indicates that the obtained responses by no means adequately reflect the marginal aspects of the concepts. Only for the characteristics *happy* and *sad* in Tribe A and *fat* and *skinny* in Tribe B is there a fair correspondence between observed and expected marginal probabilities.

There is, however, some overall correspondence between the expected proportions derived from the conditional probability-model described above and the proportions observed in the experiment. The conditional probabilities are not reflected perfectly,

TABLE IV
DIFFERENCES BETWEEN TRIBES A AND B, THEORETICAL AND OBTAINED

Stimulus-words		Responses							
		tall	short	fat	skinny	happy	sad	black	white
tall	E	.00	—	-.07	.07	.08	-.07	.00	.00
	O	-.07	—	.04	.00	.06	-.02	-.03	.02
short	E	—	.00	.20	-.20	-.25	.25	.00	.00
	O	—	-.02	.20	-.15	-.17	.14	.00	.01
fat	E	-.25	.25	.00	—	-.25	.25	.00	.00
	O	-.32	.22	-.01	—	-.24	.35	.00	.01
skinny	E	.06	-.06	—	.00	.07	-.06	.00	.00
	O	.07	-.17	—	-.10	.24	-.03	.00	.00
happy	E	.10	-.10	-.08	.08	.00	—	.00	.00
	O	.17	-.09	-.07	.11	-.23	—	.00	.12
sad	E	-.16	.16	.12	-.13	—	.00	.00	.00
	O	-.17	.20	.15	-.08	—	-.10	.00	.00
black	E	—	—	—	—	—	—	—	—
	O	.09	-.14	.00	-.05	.04	.27	-.21	—
white	E	.00	.00	.00	.00	.00	.00	—	.00
	O	-.05	.08	.08	-.05	.15	.08	—	-.30
Differences in marginal probabilities	E	0	0	0	0	0	0	0	0
	O	-.10	.10	.32	-.32	-.17	.17	.01	-.01

but there is an association. Although admittedly not an ideal index in the present situation, product-moment correlations between observed and expected proportions were computed for each tribe to obtain some estimate of the relationship. These correlations of 0.64 for Tribe A and 0.60 for Tribe B indicate some correspondence between the conditional probability model described above and response-frequencies in word-association tests.

To the extent that the conditional probabilities among a concept's referents are reflected by the word-association responses, the meanings of Tribes A and B should be differentiated. Differences between the expected probabilities for the two tribes are given in Table IV along with differences between the observed proportions. Rows labeled E in this table give the expected differences between the tribes, and rows labeled O give observed differences. A comparison of corresponding E and O rows indicates a marked agreement between the two sets of differences.

DISCUSSION

The results indicate that the word-association test used in the present experiment differentiated the meanings of Tribes A and B (see Table IV). Such a differentiation suggests that the joint probabilities for the occurrence of combinations of characteristics were reflected by the technique in some form or other. The correspondence between the expected probabilities and the observed proportions for each tribe suggests that the reflection of these joint probabilities is at least in part analogous to the description given in the conditional probability model. It must be remembered, however, that this model involved only conditional probabilities among pairs of attributes. If higher order combination has been included as well, such as the probability of occurrence for the characteristic *tall* given the occurrence of both *happy* and *fat*, the predictions might have been better.

The results, then, suggest that word-association tests show some indication of a concept's meaning and some sensitivity to differences among similar concepts. However, it must be pointed out that these meanings are reflected only in terms of simple frequency indices and that nothing approaching the measurement properties of the semantic differential is available.⁶ Furthermore, the marginal aspects of a concept's meaning were not adequately reflected by word-association, whereas they were measured almost perfectly by the semantic differential.⁷ In any experimental application of word-association as a measure of meaning, it may also prove necessary to restrict the types of possible responses elicited, and in many practical situations the desirable restrictions may be unknown.

The word-association technique used in the present experiment was restricted only in that responses opposite in meaning to the stimulus-word were ignored. Other restrictions could have been made, some of which might have improved the results. For example, when given the stimulus-word *tall*, *S* might have been asked to give the first word that came to mind from the list *fat*, *skinny*, *happy*, *sad*, *black*, and *white*. Or *S* might have been asked, "When *E* says *tall*, which word do you think of, *happy* or *fat*?" The latter highly structured procedure is analogous to the method of triads and suggests the use of multidimensional scaling methods as possible measures of meaning.⁸

SUMMARY

A word-association test was given to see which statistical aspects of the meaning of a concept it would reflect. An artificial concept, 'Tribe,' was

⁶ Messick, Metric properties of the semantic differential, *Educ. psychol. Measmt.*, 17, 1957, 200-206.

⁷ Solley and Messick, *op. cit.*, 168-169.

⁸ W. S. Torgerson, Multidimensional scaling: I. Theory and method, *Psychometrika*, 17, 1952, 401-409; Messick, *op. cit.*, 1956, 83-91.

constructed, for which the 'true' frequency-distribution of referents would be known. Two tribes were used; their referents were drawings of 'stickmen' which varied as to whether they were tall or short, fat or skinny, happy or sad, and black or white. The two tribe-concepts were identical with respect to the individual or marginal occurrence of these characteristics, but they differed with respect to the joint occurrence of combinations of characteristics. It was of interest to see whether a word-association test would differentiate the two tribes and thus reflect the joint occurrences.

The Ss were first taught the two concepts by means of a probability-learning procedure in which they were asked to guess the characteristics of a hidden stickman. One group of nine Ss learned one tribe, and another group of nine learned the other. Immediately following the probability-learning series, Ss were told that *E* would give some stimulus-words to which they should respond with whatever words occurred to them. The words *tall*, *short*, *fat*, *skinny*, *happy*, *sad*, *black*, and *white* were given by *E*, and Ss' responses were recorded. The results indicated that the word-association test did reflect the joint occurrences of various perceptual characteristics within a concept's referents, thus differentiating the meanings of the two concepts.

THE PATTERN OF ERRORS IN DELAYED RECALL OF SERIAL LEARNING AFTER INTERPOLATION

By JAMES DEESE and VICTOR J. MARDER, The Johns Hopkins University

Despite an enormous literature on negative transfer and retroactive inhibition there has been relatively little attention paid to the pattern of responses in the observed transfer effects. An earlier study from this laboratory showed that the pattern of errors in an experiment on retroactive inhibition depends upon the organization of material used in the learning tasks.¹ In serial learning by the method of anticipation, the chief effect of an interpolated list upon relearning was to increase the frequency of errors based upon items from the original list given at inappropriate positions (within-list errors). Intrusions from the interpolated list accounted for a negligible fraction of the total errors. In paired-associate learning, on the other hand, errors based upon inappropriate items from the original list were infrequent—most of the errors were between-list intrusions. It was concluded that these findings reflected the context in which individual items appeared.

At the same time it was argued that the relatively strong contextual boundaries that separated the lists in immediate recall of serially organized material would deteriorate with the passage of time between interpolated learning and recall. Thus, for serial learning, the frequency of errors at recall based upon intrusions of interpolated material would increase with increase in the delay between interpolated learning and recall. Some support for this possibility appears in an early study by McKinney and McGeoch, but a more thorough experimental investigation of this point is required since both the experimental variables and method of analysis used in the study by McKinney and McGeoch are inadequate for dealing with the problem of the influence of context upon recall.²

In order to study the pattern of responses in verbal learning, the traditional methods of presentation should be abandoned, for the familiar methods that make use of paced presentation enormously decrease the frequency of overt errors.³ Most of the errors in paced learning are simply omissions. The assumption has sometimes been made that omissions, in large part, represent truncated or implicit responses and that these show about the same patterns as the overt responses. It seems more desirable, however, to make use of a method that would allow the truncated or implicit responses to occur.⁴ In the study of paired-associate retention, other investigators have made use of a method of modified free recall; this method requires

* Received for publication February 28, 1956. This research was in part supported by funds from National Science Foundation grant NSF-61369.

¹ James Deese and G. W. Hardman, An analysis of errors in retroactive inhibition of rote verbal learning, this JOURNAL, 67, 1954, 299-307.

² Fred McKinney and J. A. McGeoch, The character and extent of transfer in retroactive inhibition: Disparate serial lists, this JOURNAL, 47, 1935, 409-423.

³ James Deese and F. H. Kresse, An analysis of errors in rote serial learning, *J. exp. Psychol.*, 43, 1952, 199-202.

⁴ A. W. Melton and J. McQ. Irwin, The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses, this JOURNAL, 54, 1941, 157-173.

that Ss be interrupted every few trials or so for a somewhat constrained free recall test.⁵ This method is less satisfactory, at least for the study of serial learning, than that of allowing Ss unlimited response time for each item during the actual learning. Consequently, in the present study the method of anticipation usually used in the study of serial learning was modified so that unlimited time for responding was allowed, and instructions for forced guessing were given the Ss. This procedure more nearly exhausts the Ss' potential for responding than any other method.

The purpose of the present investigation is to determine the pattern of errors in recall of original material as a function of delay between interpolated learning and recall. The hypothesis under test is that errors based on the interpolated list will be small immediately after interpolated learning and that they will increase as the delay between interpolated learning and recall increases. No other hypothesis about the pattern of responses is made. Thus an increase in between-list errors at recall may involve an increase in the total number of errors, or it may involve displacement of errors from other sources.

METHOD AND PROCEDURE

The Ss in this study were 80 girls from the junior and senior classes of Baltimore city high schools. These Ss were divided into 5 groups of 16 Ss each. Each S was tested individually, and no S had previously taken part in any experimentation on verbal learning. All Ss took part in a psychophysical experiment in addition to the verbal learning experiment, and those Ss whose recall for original learning was delayed until a second session were informed that the second session was for the purpose of continuing the psychophysical testing. Ss were volunteers who were paid for their work.

The items learned by Ss were adjectives taken from Melton's lists.⁶ These same items had been used in a previous study of patterns of errors in immediate recall.⁷ A practice list of six items was given to all Ss for two trials; E used this opportunity to correct S's procedure. These items were presented on flash cards.

The original list and the interpolated list were made of synonymous adjectives with a median similarity rating on Melton's scale of 2.36. Both lists consisted of 12 items; in the case of the original list the items were presented for 7 trials, and in the case of the interpolated list, 6 trials were given. These lists were presented on a memory-drum, and the drum was halted after each item until S had anticipated the next item or made an error. One recall trial was administered after varying periods of delay. The Ss were not informed that there would be a recall of the original list until it was time for the recall-trial.

The procedure for each condition of delay is outlined below. Notice that there is a traditional retroaction-control for only one of the delay periods.

Four-minute delays. The four minutes that elapsed between the end of interpolated learning and the recall of the original list were occupied by E giving a brief account to S of the next experiment in which S was to participate. This account was interrupted by E's request that S attempt to anticipate the items of the original list for one additional trial.

⁵ B. J. Underwood, 'Spontaneous recovery' of verbal associations, *J. exp. Psychol.*, 38, 1948, 429-439.

⁶ E. R. Hilgard, Methods and procedures in the study of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 548.

⁷ Deese and Hardman, *op. cit.*, 299-307.

Two-hour delays (experimental). The *Ss* were occupied during the 2-hr. delay period by participation in another experiment; in this experiment *Ss* were required to estimate the number of tones in short bursts of brief tones presented at near threshold level. This experiment was interrupted for periodic rest-intervals, during which *E* conversed with *S*.

Two-hour delays (control). No interpolated list was given to the *Ss* in the 2-hr. delay control condition. The time taken for interpolated learning in a matched experimental *S*, plus 2 hr., intervened between original learning and test of recall for each *S* in this group. The *Ss* in this group likewise participated in an experiment during the delay period.

Twenty-four hour delays. The *Ss* in the 24-hr. delay condition took part in the psychophysical experiment immediately after interpolated learning. At the end of this experiment an appointment was made for *S* to return at the same time the next day for further psychophysical testing. At the outset of the next session, one trial of recall was given.

Forty-eight hour delays. The procedure for the *Ss* in the 48-hr. condition was exactly the same as that for the 24-hr. group, except that the *Ss* were asked to return two days later.

All errors were classified into the following four types: (a) within-list errors, based upon items from the original list given in incorrect positions; (b) between-list errors, based upon intrusions of interpolated list items; (c) intrusions of items from the practice list; and (d) errors not based on any of the above categories.

RESULTS

The mean number of errors for all *Ss* in the experimental groups was 4.5. There are no significant differences in total errors as a function of delay; an analysis of covariance for total errors per *S* (corrected for regression from original learning) for all of the four experimental groups yields an *F* of 0.94 for a comparison of the mean square for delays with the mean square based upon individual variability. This finding is not altogether inconsistent with earlier experiments which find small and not always systematic differences due to delay between interpolation and recall of material like that used in the present experiment.⁸ The mean number of total errors for the 2-hr. control was 2.4. The difference between the mean number of errors for this control and the appropriate experimental group (mean of 4.2) yields a *t* of 2.02. Thus, in the case of the 2-hr. delay, there is a significant retroactive effect of interpolated learning. With the same material, Deese and Hardman found a significant retroactive effect with no delay between interpolated learning and recall, so it appears probable that retroaction occurs in all of the four experimental conditions, though a significant retroaction effect is not necessary for the hypothesis under study.⁹

Certain subclasses of errors do change significantly as a function of delay. The most important of these is the class of between-list errors. An analysis of variance

⁸ J. A. McGeoch, Studies in retroactive inhibition: II. Relationships between temporal point of interpolation, length of interval, and amount of retroactive inhibition, *J. gen. Psychol.*, 9, 1933, 44-57; Leo Postman and T. G. Alper, Retroactive inhibition as a function of the time of interpolation of the inhibitor between learning and recall, this JOURNAL, 59, 1946, 439-449.

⁹ Deese and Hardman, *op. cit.*, 299-307.

of between-list errors at recall for different delays yields an F of 3.10. This is significant at between the 1% and 5% level. Fig. 1 shows that the percentage of errors attributable to between-list errors increases from about 10% for the 4-min. delay to about 30% for the longer delays. Data from the experiment by Deese and Hardman show that about 5% of the total errors at *immediate* recall are from the interpolated list, so it seems reasonable to conclude that there is an important increase in the absolute number and percentage of between-list errors.

The only other class of errors to change significantly as a function of delay is that composed of within-list errors. Fig. 1 shows that the percentage of within-list

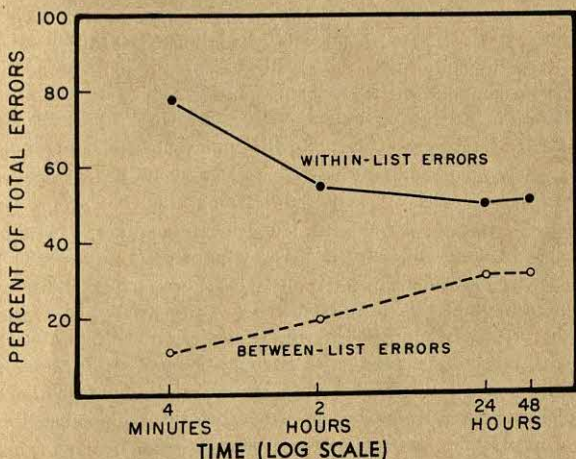


FIG. 1. PERCENTAGE OF BETWEEN-LIST AND WITHIN-LIST ERRORS AS A FUNCTION OF TIME

errors decreases with increasing delay; since the total error-frequency does not change, this reflects a decrease in the absolute number of within-list errors as well. The F for the comparison of mean square for delays with the mean square for individuals in the case of within-list errors is 2.91. This is significant beyond the 5% level.

Intrusions of responses from the practice list and other, unclassifiable errors occur in very small numbers, consequently from the present data it is impossible to know whether there are any changes in these categories. There are no responses from the practice list at a delay of 4 min. and only one or two for all S s combined at each of the other delays. There are likewise a relatively small number of unclassifiable responses at each delay, and many of these may represent clang associations, etc., with the items from the two experimental lists.

DISCUSSION

The major results of the present study can be summarized by saying that between-list errors increase and within-list errors decrease with lengthening of the delay between interpolated learning and recall. These results are consistent with the

earlier findings of McKinney and McGeoch as well as with those of Newman on the retention of serial lists.¹⁰ Furthermore, it is to be noted, the number of items from both lists used as responses by Ss during recall was approximately a constant for all delays. Thus, by the method of unpaced testing with forced guessing, individual items themselves do not appear to be forgotten, but the context in which they occur does appear to be forgotten. Items from the interpolated list displace the weaker, incorrectly located items from the original list in delayed recall. In the present experiment the interpolated items apparently do not displace items in the correct position from the original list. It is possible, though, that with difficult or more unfamiliar items or with less practice that the interpolated items could displace correct items from the original list.

A word should be said about the relation of the present results to studies in which the material to be learned is differently organized, since one of the implications of the context notion is that the pattern of errors in retroaction will depend upon the organization of the separate lists. Deese and Hardman had found that the pattern of errors was very different in the immediate recall of paired-associate and serial lists, so it might be expected that there would be a difference in the pattern of changes with delay.¹¹ In serial learning the context of each item is established by the varying strengths of association of each item with every other item in the list, while in paired-associate learning there are overwhelmingly strong associations between particular responses and the stimuli paired with them. Thus, for practical purposes, there is little context except that supplied by the stimulus associated with particular items. The retroactive effect in paired-associate learning is primarily based on the existence of two alternative responses to common stimulus-items.¹² Thus, if, as we suppose, the effect of delay is to weaken the contexts in which particular responses are to be emitted rather than to weaken the responses themselves, we would guess that errors made of responses based upon inappropriate items irrespective of list would increase in frequency relative to the frequency of responses from the interpolated list based upon appropriate items.

A study by Briggs gives us some aid on this point, since he found, in a kind of free-association type of recall test, that interpolated responses do decrease in frequency as a function of delay in paired-associate learning.¹³ This decrease of interpolated list responses at recall seems to be associated with both an increase in the frequency of appropriate original list response and inappropriate responses, so that in Briggs' study only a portion of the decrease of interpolated responses with delayed recall could be described as the result of a disappearance or weakening of context.

SUMMARY

The purpose of the present experiment was to study the effects of delaying the test of retention upon the pattern of errors in recall of a list of words serially presented following the learning of an interpolated list. All Ss practiced a list of

¹⁰ McKinney and McGeoch, *op. cit.*, 409-423; E. B. Newman, Effect of crowding of material on curves of forgetting, this JOURNAL, 52, 1939, 601-609.

¹¹ Deese and Hardman, *op. cit.*, 299-307.

¹² *Ibid.*, 299-307.

¹³ G. E. Briggs, Acquisition, extinction and recovery functions in retroactive inhibition, *J. exp. Psychol.*, 47, 1954, 285-293.

12 adjectives for 7 trials, and immediately afterwards they practiced an interpolated list of similar adjectives for 6 trials. After delays ranging from 4 min. to 48 hr., the Ss were given an anticipation-test of recall for the original list. In both practice and recall, Ss were unpaced and they were forced to guess a word for every item. The responses in recall were categorized as (1) correct, if they were in the correct position, (2) within-list errors, if they came from the original list but were given in incorrect positions, (3) between-list errors, if they came from the interpolated list, (4) practice-list errors if they came from the practice list given before the experiment, and (5) unclassifiable errors. The number of between-list errors increased as the delay between interpolated learning and recall increased, while the number of within-list errors decreased with increasing delay. The frequency and changes with delay of the other classes of errors were small. The number of items recalled from the two lists combined as well as the number of items from the original list given in correct position remained a constant as a function of delay. Consequently it was concluded that the context of items is forgotten much more readily than the items themselves.

ESTIMATIONS OF LOUDNESS BY A GROUP OF UNTRAINED OBSERVERS

By JOSEPH C. STEVENS and ENDEL TULVING, Harvard University

In a recent issue of this JOURNAL S. S. Stevens reported several experiments, the results of which support the validity of the method of magnitude-estimation in establishing a scale of loudness.¹ He relied principally upon two methods of research. In one, *E* assigned a numerical value to a standard loudness and *O* estimated other variable loudnesses in terms of it. In the second, no standard was used and *O* estimated the loudnesses of various sounds on a scale of his own choosing.

Since the results reported in Stevens' paper were in general obtained from sophisticated *Os* (graduate students and members of the staff), the question may be asked whether less sophisticated *Os* would give comparable results. In assessing the validity of the loudness-scale derived from direct estimates of magnitude with no designated standard, it might be advisable to use naïve *Os*. That is to say, if *Os* have estimated loudness in experiments in which *E* designated a standard by means of a particular number, they may come to experiments in which no standard is designated with a predisposition to use numbers similar to those previously designated. If this is the case, it would be difficult to distinguish between the two procedures.

A second question to be considered is whether the method of 'magnitude-estimation' yields results when employed with groups of *Os* that are comparable to those obtained from *Os* serving alone. Except for an early study by Ham and Parkinson, most experiments on loudness have been done with individual *Os* under rather carefully controlled conditions.² From the point of view of acoustical engineering, it would be profitable to compare the loudness-scale based on the estimates of a large group of *Os* in a relatively diffuse sound-field with the scales obtained from individual *Os* wearing earphones.

It was with this problem in mind that we decided to test the two methods of estimating loudness described above.

Procedure. The experiments were performed on 70 Harvard and Radcliffe undergraduate students in introductory psychology during the course of a class session. As far as we know, only two members of the class had ever before participated in an experiment on loudness. Because of their number and their inexperience in judging loudness, it was thought that this group of students would provide a suitable population on which to test the generality of the sone scale.

Experiment I. The *Os* were informed that they were to take part in two experi-

* Received for publication May 29, 1956. This research was carried out under Contract N5ori-76 between Harvard University and the United States Office of Naval Research (Project NR142-201, Report PNR-184).

¹ S. S. Stevens. The direct estimation of sensory magnitudes—loudness, this JOURNAL, 69, 1956, 1-25.

² L. B. Ham and J. D. Parkinson, Loudness and intensity relations, *J. acoust. Soc. Amer.*, 3, 1932, 511-534.

ments in the scaling of psychological magnitudes. For Experiment I, the following instructions were read to the class:

We are going to present to you through the loudspeaker a series of noises which will last for a few seconds each, with a short time-interval between successive noises. When you hear the first noise put down on your paper, opposite No. 1, any number at all which you think appropriate to represent the loudness of that noise. When you hear the second noise assign a number to it, beside No. 2 on your paper. Try to make the ratios between the different noises correspond to the ratios between the loudnesses of the noises. In other words try to make the number proportional to the loudness as you hear it. Thus, if you give the number X to the first noise, and the second noise sounds six times as loud to you as the first, assign the number 6 times X to it. Or, if the second noise sounds 15 times as soft as the first, assign the number X divided by 15 to it. Be sure to avoid looking at your neighbor's paper, as independence of judgment is critical to the experiment.

Special precautions were taken to avoid suggesting any particular number to designate the loudness of the first noise.

Eight intensive levels of white noise were presented for approximately 3 sec. each, with about 7 sec. intervening between presentations. The noises were delivered through a 12-in. loudspeaker in a bass-reflex cabinet placed at the front of the classroom. The bandwidth of the noise was limited by a low-pass filter having a cut-off at 10,000 $\sim$. The loudest sound presented had a sound-pressure level of approximately 105 db. (*re* 0.0002 dyne/cm.²) as measured on a General Radio sound-level meter located 3 ft. in front of the source. To produce and to measure sounds below this level we relied on an attenuator connected between the noise generator and an amplifier. The levels presented were 85, 100, 75, 55, 90, 105, 65, and 95 db., in that order. The sound-pressure levels of these noises were, of course, lower at points farther removed from the loudspeaker. A check with a sound-level meter showed that the levels were approximately 10 db. lower when measured in the middle of the classroom with the *O*s present. Since, however, the slope of the sone-scale is, as we believe, essentially invariant with sound-pressure level, the fact that our *O*s were exposed to different absolute levels should not be critical.

Each *O* recorded his numerical estimation of the loudness of each sound. The records were collected before Experiment II was begun. The results of the first experiment are discussed below.

Experiment II. For Experiment II, the following instructions were read:

In this experiment we are going to present to you pairs of noises in succession. The first of each pair is the standard, and it will stay the same throughout the experiment. The second of each pair is the variable and will differ from one pair to another. We are going to call the loudness of the standard '10.' Your task is to estimate the loudness of the variable. In other words, the question is: if the standard is called '10,' what would you call the variable? Use whatever numbers seem to you appropriate—fractions, decimals, or whole numbers. For example, if the variable sounds seven times as loud as the standard, write '70'; if the variable sounds 1/5 as loud as the standard, write '2'; if 1/20 as loud, write '1/2' or '0.5,' etc. Try not to worry about being consistent. Try to give the appropriate number to each noise regardless of what you may have called some previous one.

The sound-pressure level of the standard noise presented before each variable was 85 db., determined in the fashion described above. The levels of the variable noises were 90, 65, 105, 95, 55, 85, 100, and 75 db., in that order. The standard was in

each case presented for about 3 sec., followed by a variable of the same duration, with about 2 sec. of silence between them. Approximately 10 sec. were allowed for the *Os* to record their estimates before the next pair was presented.

Results. The estimates given in Experiment I were transformed by multiplying each *O*'s judgment of the loudness of the first noise by a factor that made the resulting product equal 10, and then multiplying each of his successive judgments by the same factor. This operation puts all the judgments of loudness on a common scale. The medians of the transformed estimates of loudness for Experiment I are represented by circles in Fig. 1. In the same figure the median estimates of loudness obtained in Experiment II (in this case actual estimates) are designated by triangles. The interquartile ranges are indicated by the length of the vertical lines.

The two straight lines in Fig. 1 have the same slope as the sone-function derived by Stevens from the pooled results of about a dozen previous experiments.³ The slope of these lines is such that an increase of 10 db. in sound-pressure level corresponds to a twofold increase in estimated loudness. In view of the general conditions of our experiment, which were certainly less well controlled than those in the laboratory investigations, the close fit between our results and the line representing the sone-function is significant. The fit is apparently a little better for the data obtained in Experiment I where no standard was designated. Previous work would lead us to expect this result. Furthermore, in Experiment II the deviations from the predicted line are in the same direction as has been found in earlier experiments, *i.e.* when the standard is set near the middle of the range, very faint sounds appear to be underestimated and very loud ones overestimated.

This systematic effect has been interpreted by Stevens essentially as follows: In making estimates of loudness relative to a standard the *O* does two things: he judges the loudness of the variable with respect to the standard, but to a small extent he also tends to report something about the absolute loudness of the variable. It is as though a given estimate were influenced by two factors: the ratio of the variable to the standard and the absolute value of the variable. How this process will affect the slope of the loudness-function has been shown to depend on the level of the standard sound. In our experiment, the placing of the standard near the level of 'medium loudness' produced the expected increases in slope near the extreme ends of the function. For a more detailed discussion of the effects of various experimental procedures upon deviations of this kind the interested reader is referred to Stevens.⁴

As shown in Fig. 1 the variabilities of the judgments under these two experimental methods are roughly comparable. The overall variability is approximately the same as in the studies employing individual *Os*.

It is of interest to examine the numerical values assigned by the *Os* to the first sound in Experiment I. Table I shows the distribution of these numbers. The number 10 was highly preferred by this group of *Os*. Other popular numbers were 20, 5, 50, and 100.

It appears that the *Os*' estimates of loudness in Experiment I depend to some

³ S. S. Stevens, The measurement of loudness, *J. acoust. Soc. Amer.* 27, 1955, 815-829.

⁴ S. S. Stevens, this JOURNAL, *op. cit.*, 6-22.

extent upon their choice of a low or high number to represent the loudness of the first sound. Evidence of this dependency is shown in Fig. 2 where curves are plotted for subgroups of the 70 *O*s divided as follows: (1) 20 *O*s who assigned

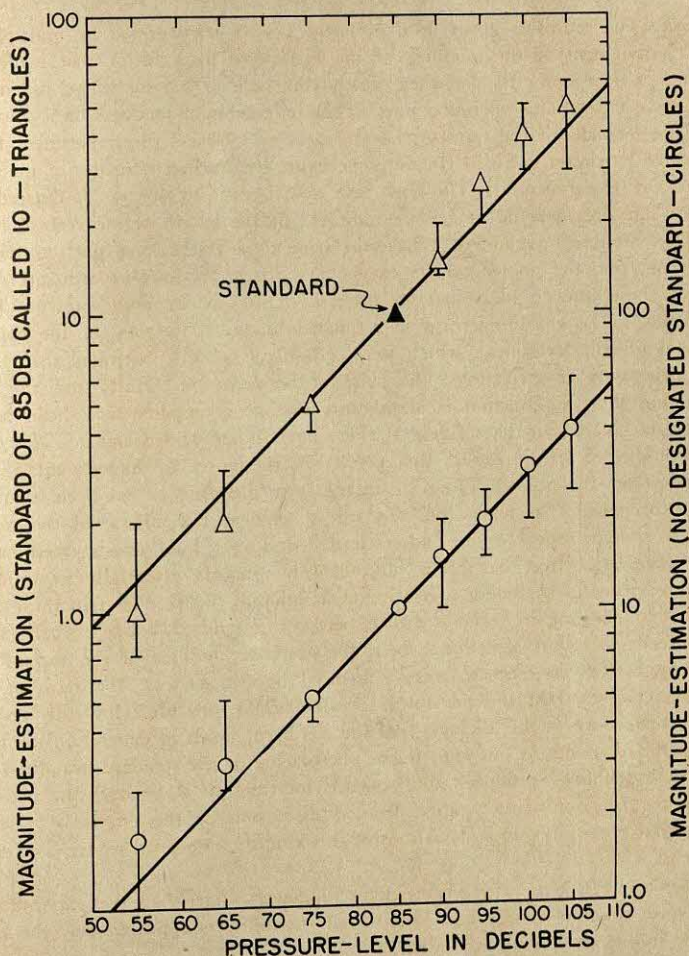


FIG. 1. MAGNITUDE-ESTIMATIONS OF THE LOUDNESS OF WHITE NOISE
Circles refer to Experiment I in which a standard was not used. Triangles refer to Experiment II in which the loudness of the standard (85 db.) was called 10. Each point is the median of the judgments of 70 *O*s. The vertical lines indicate the interquartile ranges. The straight lines show the slope of the sone-scale of loudness for a 1000~ tone.

a number smaller than 10 to the first stimulus; (2) 30 *O*s who used numbers from 10 to 30; and (3) 20 *O*s who used numbers 40 and above. The median estimates of Group 1 show the largest deviations from the predicted line at the low end of

TABLE I
FREQUENCY-DISTRIBUTION OF NUMBERS CHOSEN IN EXPERIMENT I TO
DESIGNATE THE LOUDNESS OF THE FIRST STIMULUS

Number chosen	Frequency	Number chosen	Frequency
1	3	20	7
2	1	24	1
4	4	30	1
5	5	40	3
6	3	50	7
7	2	70	2
8	2	80	1
10	18	100	6
12	2	120	1
17	1		

$N=70$

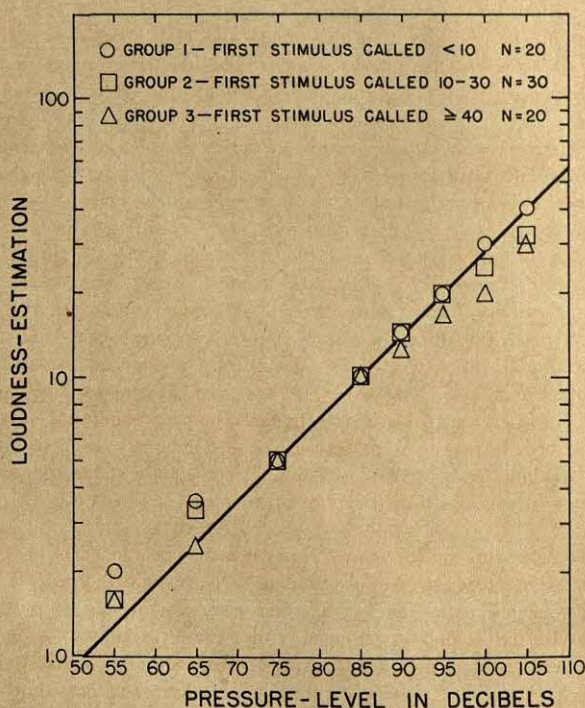


FIG. 2. EFFECT OF THE INITIALLY CHOSEN NUMBER UPON THE LOUDNESS-ESTIMATES IN EXPERIMENT I

The circles represent the median estimates of the 20 *O*s who selected a number smaller than 10 to represent the loudness of the first noise. The squares stand for the 30 *O*s whose first choice ranged from 10 to 30. The triangles stand for the 20 *O*s whose first choice was 40 and above. The straight line represents the sone-function.

the scale, whereas the medians of Group 3 deviate most markedly at the high end of the scale. Relative to the median judgments of the entire group, the *O*s who choose a small initial number tend to overestimate the fainter sounds, and the *O*s who choose a large initial number tend to underestimate the louder sounds. It would seem as though the *O*s felt some degree of constraint imposed by their relatively extreme initial choice of a number. The judgments of Group 2, who started with medium sized numbers, deviate from the line representing the sone-scale less markedly than the two other groups.

In Experiment I, in which there was no standard, the results from our unsophisticated *O*s followed a function whose slope is a little flatter than the function obtained from more experienced *O*s. It seems not unlikely that the judgments of naive *O*s would tend in this direction, for the factors that produce errors in these experiments probably work more in one direction than in the other. One such factor may be the failure of the naive *O* to make true ratio judgments. For example, he may not understand completely what is meant in the instructions: "make the numbers proportional" to something. As a result, instead of assigning numbers in proportion to the loudnesses, he may attempt to rank order them; or he may assign numbers to represent the intervals rather than the ratios between the loudnesses. When any or all of the *O*s adopt these attitudes, the effect in general is to flatten the slope of the loudness-function. Suppose, for example, *O* assigned the numbers 1 to 8 to stand in an ordinal fashion for the eight sounds presented in the first experiment. The ratio between the loudness of the loudest noise and the softest would then appear to be only 8 to 1, whereas the sone-function based on many previous experiments predicts a ratio of 32 to 1. Judging ratios is probably harder than merely ordering stimuli, and the lazy or inattentive *O* probably slips easily into an 'ordinal' attitude. It seems reasonable to expect the practiced *O* to be less subject to these pitfalls and more likely to give true ratio-judgments.

Despite these difficulties, the results of these two experiments add further evidence to the validity of the sone-scale as a measure of the typical listener's assessment of the loudness of a sound. Moreover, the fact that the slope of the loudness-function obtained from a group of listeners under classroom conditions is very similar to those obtained in previous studies with well-trained *O*s, suggests that the relation between loudness and intensity of sound is sufficiently stable to permit the scaling of loudness by unsophisticated groups of people.

Summary. Seventy inexperienced *O*s estimated the loudnesses of a series of white noises presented through a loudspeaker in a classroom. In the first experiment, no standard was designated, and each *O* estimated loudnesses on a scale of his own choosing. In the second experiment, a numerical value was assigned by *E* to a standard noise which preceded the presentation of each variable. The results of the two experiments verify in several respects the loudness-functions previously reported by S. S. Stevens. The functions are approximately linear on a log-log plot. In the experiment in which a standard was not designated, the slope was such that a twofold increase in loudness corresponds approximately to a 10-db. increase in intensity. Judgments were seen to depend somewhat upon *O*'s choice of a high or a low number to represent the loudness of the initial noise. The slope of the loudness-function was considerably steeper when a standard of medium loudness preceded the variable stimuli.

VERBAL CONTEXT AND FREE ASSOCIATION-TIME

By CHARLES N. COFER and THOMAS J. FORD, University of Maryland

The description and analysis of verbal behavior in terms of generalization-gradients based on meaningful relationships or on common mediating processes have offered some prospect of increasing our understanding of the ways in which verbal responses interact.¹ In turn, such knowledge, as it develops, will permit the analysis of other processes, such as perception, learning, and problem-solving, in which verbal factors play important roles. The present paper reports an investigation which falls in the general framework of the problem of context in verbal behavior. It is clear that in many instances the verbal response which occurs does so because of the particular circumstances that have preceded the stimulus to which it is a response. Relatively little is known either about the factors involved in context or about the processes through which they have their effects.² The hypothesis here examined is that meaningful relationships among words (*e.g.* synonymy) may operate to affect the occurrence in speech or in an experimental setting of words so related.

It seems reasonable to assume, when one sees, hears, thinks of, or speaks a specific word, that words closely related to that word (*e.g.* its synonyms) become more readily available and more likely to arise than do other words.³ This assumption is examined in the present study. The association-time to a given word was measured when the word was preceded (1) by no specific verbal context; (2) by the context of an unrelated word; and (3) by the context of a word which had some degree of synonymy to the given word.

Methods. The design of the experiment required that pairs of words be available between the members of which there was a known degree of synonymy. The procedure involved the presentation of a 'standard' word, its pronunciation, the presentation of a second word (the stimulus-word), and the giving of a free association to this second or stimulus-word.

Materials. Haagen's list was used as the source of word-pairs.⁴ Twenty-five pairs were chosen, the stimulus-words of which had Thorndike-Lorge frequency-ratings of either A or AA, so that the degree of synonymy between the stand-

* Received for publication June 4, 1956. This report is based on a master's thesis prepared by the second author, An investigation of mediated generalization among synonyms along a gradient of similarity of meaning, University of Maryland, 1952. It is also abbreviated from Technical Report No. 19, Contract N7onr-397, T. O. III, with the Office of Naval Research.

¹ C. N. Cofer and J. P. Foley, Jr., Mediated generalization and the interpretation of verbal behavior, *Psychol. Rev.*, 49, 1942, 513-540; C. E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 690-727.

² Davis Howes and C. E. Osgood, On the combination of associative probabilities in linguistic contexts, this JOURNAL, 67, 1954, 241-258.

³ Cofer and Foley, *op. cit.*, 513-540; B. F. Skinner, The verbal summator and a method for the study of latent speech, *J. Psychol.*, 2, 1936, 71-107.

⁴ C. H. Haagen, Synonymy, vividness, familiarity and associative value of 400 pairs of common adjectives, *J. Psychol.*, 27, 1949, 453-463.

ard and stimulus-words varied.⁸ The stimulus-words selected were presented tachistoscopically one at a time to 10 Ss who gave a single word in free association to each. Words retained for the experiment were those which, under these conditions, did not differ significantly from the other members of the frequency-category (A or AA) in associative reaction-time. Fourteen pairs were thus selected, in seven of which the stimulus-word had an A rating and in the other seven an AA rating. The degrees of similarity between the standard and the stimulus-words ranged from 0.9 (closely similar) to 4.8. To have pairs with no synonymity between the members, stimulus-words from each category were randomly as-

TABLE I
STANDARD WORDS, STIMULUS-WORDS, SYNONYM AND FREQUENCY RATINGS
AND MEAN ASSOCIATIVE REACTION-TIMES (IN SEC.)

Stimulus-word	Frequency rating	No context (10 Ss)	Unrelated-context (10 Ss)		Synonym-context (25 Ss)	
		Mean for stimulus-word	Standard word	Mean for stimulus word	Standard word	Mean for stimulus-word
evil	A	1.12	former	1.01	wicked	1.2
giant	A	1.17	constant	1.01	mammoth	1.4
tiny	A	0.95	sullen	0.96	little	1.8
friendly	A	1.27	empty	1.64	pleasant	2.0
hollow	A	1.04	cautious	1.85	empty	2.4
silent	A	0.88	beloved	1.00	sullen	3.9
ancient	A	0.81	sacred	1.08	former	4.5
country	AA	1.02	royal	1.16	rural	0.9
clouded	AA	1.32	double	0.98	unclear	1.8
talking	AA	1.32	rural	0.94	spoken	2.3
flowing	AA	1.11	unclear	1.49	liquid	2.7
burning	AA	1.07	spoken	1.00	ardent	3.4
open	AA	0.90	urgent	1.18	artless	4.7
crying	AA	0.85	ardent	1.12	urgent	4.8

signed to standard words from the original pairs; and the resulting 14 pairs showed no apparent intra-pair semantic relationships. The pairs are presented in Table I. Three additional unrelated word-pairs were used as practice-materials.

Subjects. Forty-five Ss (men and women) from undergraduate classes in psychology were used. Ten, Group I, gave associative reactions to the stimulus-words alone; 10, Group II, to the stimulus-words preceded by unrelated standard words; and 25, Group III, to stimulus-words preceded by a standard word of some degree of synonymity to the stimulus-word.

Procedure. S sat before an exposure-device. After a 'ready' signal, E presented every word singly and S either pronounced it (standard word) or gave an association to it (stimulus-word). The words were printed on white cards. The standard words were underlined, the stimulus-words were not. S was instructed to pronounce the underlined words and to give free associates to the other words. Both reaction-times (reading and associative) were recorded by means of a voice-key and chronoscope. The general instruction so disguised the nature of the study that the Ss thought the experiment was directed upon the pitches of their voices as they read or gave associative responses to the words. The Ss of Group I, from

⁸ E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, 1944.

whom associative reactions to the stimulus-words alone were obtained, did not see the standard words.

The order of the pairs of words was randomized within each frequency (A and AA). For successive Ss, the A pairs and the AA pairs were alternated. The Ss saw a given stimulus-word only once.

Results. Table I gives the mean associative reaction-times for every stimulus-word for each condition of the experiment. It is clear that the reaction-times do not show a close relationship to the degree of synonymy to the standard words in the case of the A or the AA frequencies. Actually only one mean difference in each frequency-category yielded a significant *t*-value in the direction expected by hypothesis.

Further analyses may be made among the contextual conditions. While the analysis of mean differences between the synonym-context and unrelated-context and between the synonym-context and the no-context shows few significant differences, it is noteworthy that in 11 of the 14 comparisons between the no-context and the synonym-context, the associative reaction-time is longer for the latter condition. This result is directly in conflict with the hypothesis which predicts the opposite outcome. A sign-test of these comparisons yields a *p*-value at the 0.029 level (one-tailed). While not significant, 9 of the 14 comparisons show lower means for the no-context group than for the unrelated-context group. The unrelated-context condition is superior to the synonym-context group in 7 of the 14 mean comparisons.

This evidence suggests that, under the present conditions, context acts to delay the occurrence of an associative response rather than to facilitate it as had been originally hypothesized and, while the differences are small, it appears that the synonym-context may engender more delay than the unrelated-context.

Further thought concerning the task of the Ss in the present experiment suggests the possibility that our expectation was not justified to begin with. While the stimulus-word may become more available when it follows a close synonym, it may be that its most likely free associative response may often fall outside the category of meaning in which the synonymy of the standard and stimulus-words resides. Should this be the case, then one might expect that if the context does strengthen this meaning category, the associative response might be delayed through conflict with the ordinarily dominant associative response. While this effect should not be so great for the unrelated pairs, inspection of the standard words in Column 6 of Table I shows that some of them may have relations to their stimulus-words (Column 3), although not directly synonymous with them, which might also engender conflict.

These considerations suggested that study of the free associative responses to the standard and stimulus-words with reference to whatever meaningful relations obtained in any pair would perhaps clarify the nature of the results. As the associative responses were not recorded in the present study, a free association-test was made up, including all of the words used here, and data were collected from 252 Ss. This material will be reported separately, but some comments can be made here relative to the questions raised above.

With regard to the neutral-context condition, the associations obtained to the standard words in no case include either the stimulus-word paired with the

standard word or any of the associations to the stimulus-words (with any frequency). It must be concluded that so far as synonymy or commonness of meaning is concerned the neutral-context condition did provide an unrelated context. With the synonym-context pairs, however, a considerably different situation appears in the associations. Characteristically, the standard word produces the stimulus-word as a high-frequency associate, or the stimulus-word yields the standard word as a high-frequency associate, or both standard and stimulus-words produce the other as a high-frequency associate, or both of them produce in common other words as high-frequency associates. These relationships are true of 10 of the 14 pairs. In the other four pairs (*former-ancient*, *ardent-burning*, *artless-open*, *urgent-crying*) these relations are almost absent. *Old*, the high-frequency associate to *ancient*, does occasionally occur to *former*, and *burning* occasionally occurs to *ardent*; however, the associates to *artless* and *open* and to *urgent* and *crying* show few resemblances within the respective pairs. It is noteworthy that these four pairs are among the five with the least rated synonymy; the other pair, *sullen-silent*, has the least resemblance in the associations to its words of any of the remaining pairs.

From these data, then, it would appear that on the whole the synonym context was effective in the sense that the associations to the standard words tend to involve the stimulus-words and conversely.

A further finding of some interest is that the rank-order of the mean associative times for the stimulus-words under the synonym-context condition approximates the order for the same words in the no-context condition. A rho of 0.63 is obtained between these two rank orders, but the rhos between the rank orders of the neutral-context mean values and those for the no-context condition and between the means for the synonym-context and neutral-context conditions are low, being -0.19 and 0.06 respectively. These relationships, together with the associative data and the mean differences, suggest that the synonym context produced some kind of interference condition which affected most of the stimulus-words which followed synonyms in roughly the same way. It was not possible to find any commonality in the four stimulus-words (*silent*, *hollow*, *giant*, *talking*) whose mean-associative-time ranks following synonyms deviate from their ranks in the no-context condition, although they were carefully studied with respect to their various associative relationships. It may likewise be noted that consistent differences could not be observed among those neutral-context stimulus-words which did and did not deviate from their ranks under the no-context condition.

The findings reported indicate that under the synonym-context the giving of an association to a stimulus-word appears to be retarded, as compared to a no-context condition. One could interpret this as perhaps representing a tendency to avoid the content area of the synonym-relationship, an interpretation in line with a notion suggested by Lepley.⁶ On the present results, however, one can have little confidence in any interpretation, because the differences in all cases were small and because the neutral-context condition does not produce results seemingly consistent with this or any other inference. It is possible that the use of high frequency (AA and A) stimulus-words in this experiment has masked effects which might have appeared with words of lower frequencies.

⁶ W. M. Lepley, An hypothesis concerning the generation and use of synonyms, *J. exp. Psychol.*, 40, 1950, 527-530.

SUMMARY

Associative reaction-times were obtained to stimulus-words under three conditions: no verbal context, the context of a semantically unrelated word, and the context of a word having some degree of synonymy to the stimulus-word. Insofar as the small differences obtained among these conditions are dependable, the results permit the conclusion that the synonym-context produces a delay in associative response as compared with the no-context condition, a conclusion completely at variance with the results predicted. Associative data, collected independently, were seen to confirm the judgment that the synonyms were in fact related to the stimulus-words and that the unrelated words were not. The hypothesis is entertained that an interference-factor of some kind is associated with the synonym-context, thus producing the delay in associations under the synonym-context.

THE PERCEPTION OF DEPTH IN THE ABSENCE OF TEXTURE-GRADIENT

By SIDNEY WEINSTEIN, New York University

The accurate perception of depth has been traced to various binocular and monocular cues. Gibson has attempted recently to interpret the traditional monocular cues in a new theory which states that the perception of depth is impossible in the absence of a "retinal gradient of texture."¹ The validity of that assumption may be assessed by means of a photographic test of ability to judge distance, which Gibson himself has designed.² The test consists of photographs of a large open field, showing stakes driven into the ground at various distances from S. The land is flat, with a coarse, evenly structured surface providing a good gradient of texture; and in the background, mountains, foliage, and sky can be seen. In a previous study with these photographs, the accuracy of judgment was found to be unaffected by the elimination of the texture-gradient,³ but Gibson has criticized the work for having permitted the establishment of a "mental gradient."⁴ This criticism is based on the fact that, while texture was inked out under one condition, the background of mountains and sky remained, and on the fact that half the Ss took the nongradient test after the gradient one. The present experiment avoids these criticisms.

Subjects. The Ss were 60 undergraduate students of elementary psychology, naïve with regard to the purpose of the experiment. They were divided into three groups: (a) 30 Ss were tested with the gradient photographs; (b) 15 were tested with the nongradient set, and (c) the remaining 15 were tested with the nongradient photographs, but (unlike the other groups) were not given any practice-trials preceding the testing.

Tests: (a) Gradient. The test which retained the natural gradient consists of 20 achromatic photographs of a large open field.⁵ The land is flat, providing good textural gradients for distance, and in the background there are mountains, foliage, and sky. Driven into the ground in an arc at a distance of 14 yards from the camera were 15 stakes, ranging in height from 27 (No. 1) to 83 in. (No. 15), in 4-in. steps. The widths varied randomly from 2-4 in. The standards stakes were 20 in.

* Received for publication March 21, 1957. From the Bellevue Medical Center. The program of research of which this study is a part is supported by the Commonwealth Fund of New York.

¹ J. J. Gibson, *Motion Picture Testing and Research*, 1947, 189, 191, 199; Studying perceptual phenomena, in T. G. Andrews (ed.), *Methods of Psychology*, 1948, 158-188; *The Perception of the Visual World*, 1950, 67, 70, 72, 78; The perception of visual surfaces, this JOURNAL, 63, 1950, 369.

² Gibson, *op. cit.*, 1947, 189 ff. This test, known as Distance Estimation Test CP212 A, was produced by the Air Force under the Aviation Psychology Program.

³ Sidney Weinstein, *An Analysis of the Effects of Texture-Gradient, Position in the Field, and Apparent Size, on the Perception of Depth*, Unpublished M.A. thesis, New York University, 1950, 1-41.

⁴ Personal communication.

⁵ The author is indebted to Professor Gibson for making the original of this test available for photographic reproduction.

apart, except for a gap of 8 ft. in the center of the arc, which permitted *S* to see the target-stake to be judged. Four target-stakes were employed: 63, 67, 71, and 75 in. in height (No. 10, 11, 12, and 13, respectively). Each of these target-stakes was used once at each of five distances—28, 56, 112, 224, and 448 yd. Thus, there were 20 photographs, representing each combination of the 4 target-stakes at the 5 distances.

(b) *Nongradient*. The unretouched set of glossy 8 × 10 in. photographs described above comprised the *gradient* set. In the earlier study, a photographic copy of this set was modified to delete the texture. This was painstakingly done by covering, with

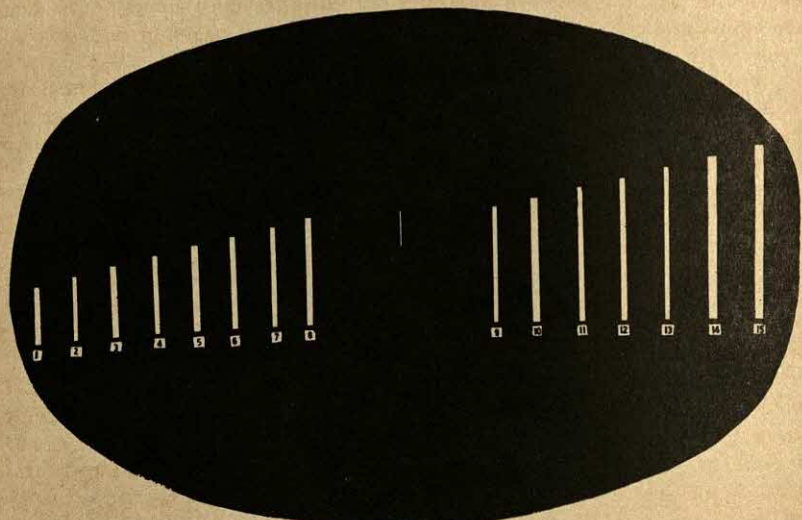


FIG. 1. COPY OF ONE NONGRADIENT FIGURE, SHOWING TARGET-STAKE IN BACKGROUND AND COMPARISON-STAKES IN FOREGROUND

In the identical gradient photograph, the base of the mountains is approximately at the level of the center of the target-stake, and the mountain top approximately at the level of the top of the target-stake. Above the mountains, sky can be seen. Below the base of the mountain is the textured foreground. In this drawing, the stake was at 56-yd. distance, and the target-stake was equal to No. 13 in height.

dense black India ink, every part of the photographed field, using a high-power magnifying lens. All that remained in each altered photograph were the 15 comparison-stakes, the single target-stake, and the background of hills and sky. In the present study, further precautions were taken to eliminate the possibility of what Gibson terms a "mental gradient." With the aid of a high-power magnifying lens, the remaining 'natural' portions of the photograph, *i.e.* the mountains in the background and the sky, were all meticulously covered with dense black India ink. The only uncovered portions of the original photographs were the 15 numbered comparison-stakes in the foreground, and the single target-stake (Fig. 1). These photographs comprised the nongradient set.

Procedure. The *Ss* of each group were tested in classrooms under normal illumi-

nation. The field conditions depicted in the photographs were described, and *S* was asked to choose the stake in the comparison-array which he considered equal to the target stake in height. He was instructed to disregard differences in the width of the stakes. Before the test proper, *Ss* in the 'gradient' group and in the first 'nongradient' group were shown three randomly chosen photographs from the set which was to follow. They were given the correct answer after each practice-trial, and were at no time aware that the practice-items were part of the actual test. The second 'nongradient' group was not allowed any practice. The 20 photographs were then presented in a randomized order for stake-size and distance-combinations. This randomized sequence remained constant for all *Ss* in the three groups. At no time did any *Ss* in the gradient group see the nongradient photographs, or *vice versa*.

Scoring. The stakes were numbered from 1 to 15 in order of height. Each response was scored for deviation. Thus, if the target-stake was No. 10, and *S* chose 13, his score would be 3. The average error (*AE*) for each photograph within each group was the arithmetic mean of these deviations.

Results. The mixed analysis of variance evaluated the effects of three variables—(a) presence or absence of the gradient, (b) the stake-size, (c) distance of target-stake—and their interactions. It was found that the presence or absence of the gradient produced no significant difference with regard to the accuracy of judgment ($F < 1$). Distance proved to be a significant variable ($F = 10.61$, $p < 0.001$), inversely related, however, to accuracy.

It can be seen that the mean *AE* increased with distance (Fig. 2). The results of *t*-tests between means of distances revealed that 448 yd. produced significantly greater errors than any other distance (*ts* ranged from 2.39 to 5.94, *ps* were < 0.05 for the two greater distances and < 0.001 for the two lesser distances). *AEs* for 224 yd. were significantly greater than those for 56 yd. ($t = 2.65$, $p < 0.05$) and for 28 yd. ($t = 3.55$, $p < 0.01$).

The results of *t*-tests evaluating the significant triple interaction ($F = 2.32$, $p < 0.05$) revealed, for the most part, no significant differences between the gradient and nongradient groups for combinations of distance and stake-size. Two differences were, however, significant; there were significantly smaller *AEs* for the nongradient than for the gradient condition at 224 yd. with Stake 13 ($t = 2.63$, $p < 0.02$) and at 448 yd. for Stake 11 ($t = 4.13$, $p < 0.001$).

The two factors previously invoked as producers of "mental gradient" were retention in the nongradient figures, of mountains and sky as background, and the prior viewing of the true gradient in a single alternation design. A third factor which might have created a "mental gradient" was practice. Since all *Ss* in the first nongradient group had been given three practice-trials in which they had been told the correct answers, such criticism seems warranted. The nongradient set was therefore presented to a new group of 15 *Ss* who were tested in the same manner, except that they were allowed no practice. Their data were then compared with those of the other two groups. The accuracy of this unpracticed nongradient group was somewhat less than that of the others (Fig. 2); however, the relation of *AE* to distance is linear, approximately parallel to those of the other groups, and reaches 'chance' accuracy only at 448 yd. An analysis of variance among the three groups for mean

AE, disregarding the variables of stake-size and distance, yielded a nonsignificant *F* (1.18).

To determine whether there were differential proportions of 'accurate' to 'inaccurate' *Ss* among the three groups, X^2 -tests were performed between groups for *Ss* with *AEs* of less than, and greater than 2.0. This was done at each distance, and for the total. None of the X^2 's reached significance, demonstrating that no group had a greater proportion of accurate *Ss* than any other in general, or at any distance (X^2 's ranged from 0.08 to 3.43, all $ps > 0.10$).

Discussion. The results indicate rather clearly that deletion of the texture-gradient in the present photographic test of depth-perception had no significant effect upon accuracy of size-judgment. The only factor which influenced accuracy differentially

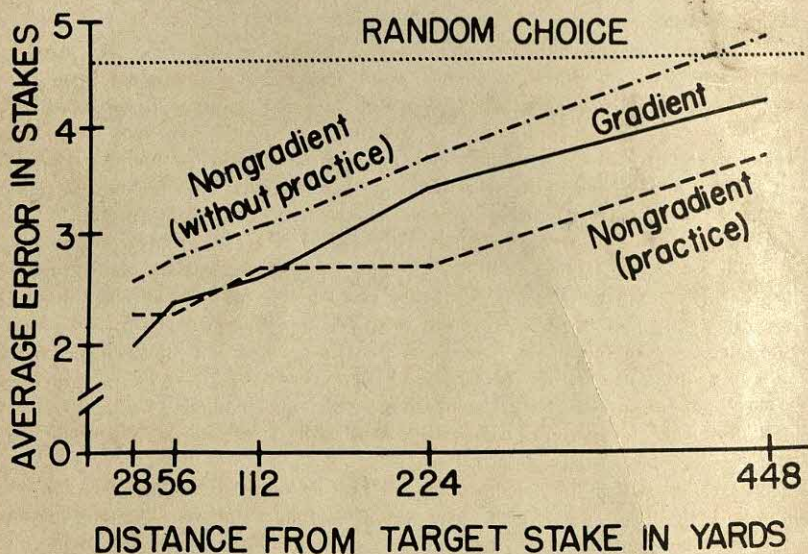


FIG. 2. MEAN *AE* FOR GRADIENT, NONGRADIENT (PRACTICE), AND NONGRADIENT (WITHOUT PRACTICE) GROUPS FOR EACH DISTANCE

was the distance of the target-stake. These results are in complete agreement with those of the earlier study and rule out an interpretation in terms of a "mental gradient." If the nongradient condition indeed created a condition "like that of the moon in the sky," with "no size constancy whatever,"⁶ Fig. 2 would show horizontal lines for *AE* as a function of distance, for both of the nongradient conditions. Moreover, both functions would overlap the chance curve. These characteristics were not obtained; furthermore, neither nongradient group produced significantly greater *AEs* than the gradient group. The slightly (although nonsignificantly) greater errors of the nongradient group without prior practice probably resulted from lack of

⁶ Gibson, *op. cit.*, 1947, 199.

practice. The *AE* function of this group, however, was linear, and parallel to those of the other groups.

Since Gibson's theory appears quite reasonable, one might indeed wonder why the perception of depth was possible in the present experiment where the texture-gradient was totally absent. There appear to be two possible reasons for the observed accuracy. The first is the remaining cue of "position-in-the-field," and the second might be termed "approximate ratio of objects of known ratio."

Position-in-the-field is the familiar classical cue which is present when a distant object appears *above* closer objects in the visual field. This situation occurs whenever an observer is at a vantage point above the plane connecting the bases of distant and close objects. In the present case the camera was above the plane connecting the bases of the target- and comparison-stakes. Differential height in the field, with all other factors constant, is a function of the distance of the target-stake. This cue is quite obvious to the observer.

The second cue which may have increased accuracy of judgment was that of approximate ratio of objects of known ratio. This cue was dependent upon the ratio of stake-width to height. *S* had been told *not* to use width as a cue; moreover the stake-widths varied randomly from 2.4 in. The cue did, however, operate in the following manner. If the retinal heights of two stakes are equal, and one appears extremely thin, and high in the field, with the other somewhat broader, and relatively low in the field, the first will be judged to be larger and at a greater distance.

In the present circumstances, the possible range of ratios of width to height was 0.03-0.06 for all the target-stakes. For example, the range of possible ratios for the smallest stake (No. 1) was 0.07-0.15. The ratios of the target-stakes and those of comparison-stake No. 1 do not overlap, with the consequence that this stake was almost never chosen as identical to any of the target-stakes. The approximate mean ratio of width to height for all target-stakes was 0.04 (*i.e.* 2/63-4/75), whereas the analogous mean ratio for all comparison stakes smaller in size was 0.09. Thus, the approximate mean width-height ratio of the target-stakes was less than half the ratio of all the stakes smaller in actual size.

It thus appears that the combined use of the cues of position-in-the-field and of approximate ratio of objects of known ratio (*i.e.* comparison-stakes) was sufficient for the adequate perception of depth and for size-estimation, even in the absence of a gradient of texture.

Summary. Depth-perception, as measured by Gibson's photographic test of size-estimation, was unimpaired when the test was presented after the elimination of the gradient of texture. It was concluded that, in the absence of texture, other cues may serve for perception of depth.

VISUAL AND KINESTHETIC COMPONENTS OF REACTIVE INHIBITION

By CARL P. DUNCAN, Northwestern University

It was recently suggested that the process of reactive inhibition, particularly as it appears in perceptual-motor learning, and the process of neural satiation in perception, are so similar that they may be indistinguishable.¹ Among other things, it was pointed out that both processes show up most clearly following a period of several minutes of continuous stimulation. In studies of neural satiation the stimulation is restricted to one afferent system. In contrast, studies of reactive inhibition, involving tasks of a perceptual-motor character, commonly provide considerable stimulation of at least two afferent systems, usually vision and kinesthetics. Thus, during a period of several minutes of massed practice on the pursuit-rotor, *S* receives continuous visual stimulation from the rotating target and turntable, and continuous kinesthetic stimulation from the circling movements of his hand and arm. It is known that under some conditions of fixation a dot that is rotated continuously will, after a few minutes, show apparent visual displacement away from the center of rotation.² Thus the reactive inhibition generated during massed practice on the rotor may be an aggregate arising from both visual and kinesthetic stimulation. This hypothesis is tested in the present study by measuring the effect on performance of each type of stimulation separately.

Method: Apparatus. The apparatus was a pursuit rotor with a bakelite turntable 6½ in. in radius. The target was a silver disk, ¾ in. in diameter, set flush with the surface of the turntable; the center of the target was 3¼ in. from the center of the turntable. The surface of the turntable was black, hence, contrasted sharply with the silver target. The turntable rotated at 60 r.p.m.; the speed was checked from time to time with a stopwatch.

The handle of the stylus consisted of the plastic handle of a screwdriver to which was hinged a brass rod 8 in. long and ⅛ in. in diameter. The hinge prevented *S* from pressing down on the rod. The rod, bent at a right angle 1¾ in. from its free end, had a silver tip and, except for the tip, was wrapped with electrician's tape. The weight of the rod exerted a force of 11 gm. on the target. When in contact, the stylus and target completed a circuit to Standard Electric clocks which recorded time on target in units of 0.01 sec.

On all trials with the *Ss* the turntable ran for 10 sec. Between successive trials the *Ss* had a 20-sec. rest-interval. The work-rest cycle was provided automatically by a control unit and was checked occasionally with a stopwatch.

* Received for publication November 26, 1956. This investigation was supported in part by Research Grant B-792 from the National Institute of Neurological Diseases and Blindness of the National Institutes of Health. Thanks are due Mr. Bruce O. Bergum for assistance in the study.

¹C. P. Duncan, On the similarity between reactive inhibition and neural satiation, this JOURNAL, 69, 1956, 227-235.

²Duncan, Development of neural satiation during judgments of diameter (Submitted for publication).

Additional apparatus, use of which will be described later, consisted of another stylus and a circular alley track made by mounting concentric circles of $\frac{1}{2}$ -in. plywood on a sheet of bakelite. The track was the same width as the diameter of the rotor-target ($\frac{3}{4}$ in.) and had the same radius ($3\frac{1}{4}$ in., measured from the center of the track) as the circle described by the rotor-target. A shield was mounted above the track in such a way that *S* was not hindered in pushing the stylus around in the track, but was prevented from seeing his hand, the stylus, or the track.

Subjects. The *Ss* were 120 men, undergraduates at Northwestern University, divided into three groups of 40 each. The groups were matched on total score over the first three trials.

Procedure. Following brief, preliminary instructions on the task, each *S* was given three trials. Then the rotor was stopped and *S* was told to put down the stylus. The *Ss* assigned to the Visual Group were then told that *E* wished to find out whether watching the rotating target was helpful. Each *S* was instructed that he would continue to stand and, upon signal, would watch the target rotating, concentrating on the size of circle the target described in order to get it firmly in mind. He was told not to stare at the target without blinking but to follow it around with his eyes in normal fashion, making sure he followed it at all times. He was told that he would do this for 5 min. and then there would be more actual practice.³ Finally *E* pointed out that since it might be difficult to concentrate on watching the target for 5 min., *E* would occasionally say "Keep watching." Following these instructions the rotor was set for continuous run and *S* watched the target for 5 min.; about every 30 sec. *E* reminded *S* to keep watching.

The *Ss* assigned to the Motor Group were told (after completing the third practice trial) that *E* wished to know whether practice at making the correct circular movement was helpful. Each *S* was instructed that he should continue standing and that, upon signal, he should pick up the other stylus and move it around in the circular alley track. (This track was mounted beside the pursuit-rotor.) He was told that the track was the same diameter as the circle described by the rotor-target and was the same width as the target, so he should concentrate on the movement and get it firmly in mind. He was told that he must move the stylus around the track once every second (same speed as the rotor), that a metronome (standing nearby) would tick every second, and that he must be at the same place in the track every time the metronome ticked.⁴ After doing this for 5 min., *S* was told, there would be further actual practice on the rotor. *E* also said that he might occasionally warn *S* to keep in time with the metronome. Then *E* started the metronome and *S* moved the stylus around the track for 5 min. From time to time *E* said, "You are doing O.K." It was rarely necessary to warn *S* to keep in time with the metronome.

The *Ss* in the Control Group rested for 5 min. after completing the third practice trial. They were merely told to sit down and rest for 5 min., after which they

³ See also R. B. Ammons, Effects of pre-practice activities on rotary pursuit performance, *J. exp. Psychol.*, 41, 1951, 187-191. Ammons found no effect of 2 min. of visual practice, so 5 min. was used in the present study to permit a more thorough test of the effects of visual practice. Use of a different type of visual practice is described in J. A. Adams, A source of decrement in psychomotor performance, *ibid.*, 49, 1955, 390-394.

⁴ A different type of motor practice was used by Ammons, *loc. cit.*

would return to practice. During the rest *S* was permitted to converse with *E*, but not about the experiment.

Near the end of the 5-min. period of interpolated activity or rest, all *Ss* were warned that they would soon return to practice on the rotor. Practice was started immediately after the 5-min. period was up. All *Ss* were then given 15 trials (10 sec. work—20 sec. rest per trial).

Throughout the experiment, time on target for each trial was recorded to the nearest 0.01 sec.

Results. The performance curves for the three groups are shown in Fig. 1. The break in the curves between Trials 3 and 4 indicates the interpolated period of 5 min. of activity (Visual and Motor Groups) or rest (Control Group). By inspection, it appears that the performance of the Motor Group is depressed below that of the Control Group, but the Visual Group is not.

Performance of each of the two experimental groups was compared with the

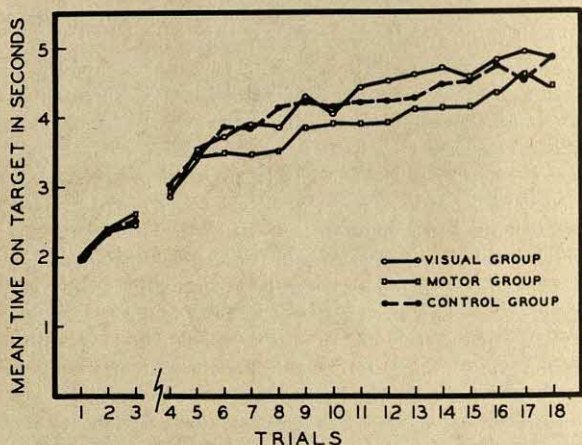


FIG. 1. PERFORMANCE OF ALL GROUPS ON ALL TRIALS

Control Group over Trials 4-18 by separate analyses of variance. Since the groups had been matched (on total score over the first three trials), Levels (paired *Ss*) was treated as a third variable, in addition to Methods and Trials, in the analyses. Table I shows the comparison of the Motor Group with the Control Group.

It can be seen in Table I that the term for Methods is highly significant when tested against the triple interaction as error. Since the Methods $\times$ Levels interaction is also significant, a more conservative test for a difference between methods can, however, be made by using this interaction as error. This test yields $F = 8.02$, which, with 1 and 39 *df*, is also significant at the 1% level. Thus, the work involved in pushing a stylus around a circular track significantly depressed later performance on the pursuit learning task.

The corresponding analysis comparing the Visual Group with the Control Group

is not shown. The term for Methods was not significant even when tested against the triple interaction ($F = 2.32$). There was some evidence that the difference between methods was not completely ineffective; Methods $\times$ Levels was significant at the 1% level ($F = 11.04$). On the whole, however, it can be concluded that the 5-min. period of watching the rotating target did not produce enough figural distortion to affect average performance.

In both of the analyses of variance the terms for Trials and for Levels were highly significant, as would be expected; Trials $\times$ Levels and Methods $\times$ Trials were not significant (all terms tested against triple interaction as error).

The results seem clear-cut. The typical depression of performance that occurs under massed practice on the rotor appears to be largely the result of massed

TABLE I
COMPARISON OF THE MOTOR GROUP AND THE CONTROL GROUP OVER TRIALS 4-18

Source	df	Mean Square	F
Methods (M)	1	24.78	12.39*
Trials (T)	14	16.36	8.18*
Levels (L)	39	30.27	15.14*
M $\times$ T	14	.69	
M $\times$ L	39	3.09	1.55†
T $\times$ L	546	.59	
M $\times$ T $\times$ L	546	2.00	

* Significant at the 1% level. † Significant at the 5% level.

kinesthetic stimulation. Even though a rotating dot will, under some conditions, show apparent visual displacement, this effect is not great enough to affect performance on the rotor.⁵ These conclusions are not necessarily vitiated by the fact that massed practice on the rotor usually produces a more severe depression of performance than that shown by the present Motor Group. Kinesthetic stimulation is probably more intense in actual practice than in merely making circular movements, since it is also necessary to make corrective movements when actually pursuing the target. Furthermore, Adams has shown that the task involved in visual discrimination of 'stylus on' and 'stylus off' the target also contributes to reactive inhibition, and this type of discrimination occurs only during actual practice.⁶ The visual stimulation experienced by the Visual Group seems, however, as intense as such stimulation would be in actual practice.

Summary. An attempt was made to test separately the effect on performance of visual and of kinesthetic stimulation which ordinarily occur together in massed practice on the pursuit-rotor. It was found that the continuous kinesthetic stimulation produced by a period of pushing a stylus around in a circular track significantly depressed later performance on the rotor, but the continuous visual stimulation produced by a period of watching the rotating target had no effect. Reactive inhibition from massed practice on the rotor seems to be entirely the result of kinesthetic stimulation.

⁵ Duncan, *op. cit.*, Footnote 20.

⁶ Adams, *op. cit.*, 390-394.

THE EFFECT OF ASSOCIATIVE VALUE AND DIFFERENTIAL MOTIVATING INSTRUCTIONS ON SERIAL LEARNING

By IRWIN G. SARASON, University of Washington

This experiment was designed to study two variables: (1) the association-value of nonsense-syllables, and (2) differential motivating instructions. The purpose of the study was to determine the effect of these variables and their interaction on performance in a verbal learning situation.

Recent work on anxiety and learning has pointed up the importance of task-difficulty,¹ the definition of which presents many methodological problems.² One way in which task-difficulty may be defined in verbal learning situations is in terms of associative values of nonsense-syllables, such as are found in tables compiled by Glaze and Hull.³ In an early study, McGeoch concluded that ease of learning was directly related to associative value,⁴ but, despite the frequent use of lists of associative values, there has been no further systematic study of the relationship. In the present experiment, three levels of associative value were compared in terms of the level of performance attained by Ss in a fixed number of trials. The second variable studied in this experiment was that of the motivating instructions under which Ss performed. Previous research has demonstrated that high and low motivating instructions have opposite effects for high as compared with low and middle-anxious Ss (anxiety defined by score on the Taylor Anxiety Scale), and at present it seems likely that highly anxious Ss differ from all other Ss in the distribution of anxiety-scores.⁵ For this reason, it seemed desirable in the present experiment to study the interaction of associative value and instructions in more representative Ss. Thus, anxiety was a controlled rather than a manipulated variable. The use of moderately anxious Ss also made possible an attempt to replicate approximately previous findings of a beneficial effect of motivating instructions for moderately anxious Ss.

Subjects. The 90 Ss of this experiment were drawn from the middle 20% of the A-scale distribution for 530 students in introductory psychology at Indiana University. The median for the distribution was 16, and the moderately anxious Ss had Taylor scores ranging from 14-18. No S with a T-score above 70 on the K-scale was used.

* Received for publication January 3, 1957. This research was performed at Indiana University.

¹I. E. Farber, The role of motivation in verbal learning and performance, *Psychol. Bull.*, 52, 1955, 311-327; I. G. Sarason, Effect of anxiety, motivational instructions, and failure on serial learning, *J. exp. Psychol.*, 51, 1956, 253-260.

²E. K. Montague, The role of anxiety in serial rote learning, *J. exp. Psychol.*, 45, 1953, 91-96; C. E. Noble, The role of stimulus meaning (*m*) in serial verbal learning, *J. exp. Psychol.*, 43, 1952, 437-446.

³J. A. Glaze, The association of nonsense syllables, *J. genet. Psychol.*, 35, 1928, 255-267; C. L. Hull, The meaningfulness of 320 selected nonsense syllables, this JOURNAL, 45, 1933, 730-734.

⁴J. A. McGeoch, The influence of associative value upon the difficulty of nonsense-syllable lists, *J. genet. Psychol.*, 37, 1930, 420-430.

⁵Sarason, *op. cit.*, 253-260; The effect of anxiety and two kinds of motivating instructions on serial learning, *J. abnorm. soc. Psychol.*, 54, 1957, 166-171.

Method. With a Hull-type memory drum, lists of 17 nonsense-syllables were presented with a 2-sec. exposure per syllable and a 6-sec. intertrial interval. The method of anticipation was employed, and three experimental lists were used. One list of 17 syllables contained syllables drawn at random from Glaze's list of syllables of 27% associative value, and the other two lists were similarly selected from the lists of 53% and 80% associative value. An examination of the lists used in this experiment reveals approximately the same proportions of every vowel in each. All but two of the syllables on the list of 80% associative-value began with different consonants and all of the 53% syllables began with different consonants. On the 27% list, three consonants were repeated once each, indicating greater intralist similarity for this list than for the others employed. All Ss were initially given 15 trials on a preliminary list, List A, which contained 6 syllables of 27%, 6 syllables of 80%, and 5 syllables of 53% associative-value. No syllable appearing in one of the three experimental lists was used in the preliminary list. The preliminary instructions were identical to those described earlier.⁶ Following the 15 trials on List A, Ss were assigned to one of nine groups, representing all possible combinations of the three levels of associative value and the three levels of instruction.

One kind of motivating instruction was called *subject-oriented* because it was intended to orient S to the importance of doing well on the experimental list, List B, in order to maintain his self-esteem. The Ss were told that the next list would be a short-form intelligence test, and that each S's score would be compared with the scores of his peers.⁷ The second kind of motivating instructions studied were called *experimenter-oriented*, and the emphasis here was placed on the 'tight spot' in which E found himself: "I know it isn't easy learning these syllables. This experiment is part of my thesis and getting my degree depends on getting good results in this experiment. I would appreciate it a lot if you would do your best to memorize these syllables as soon as possible." Control groups which received neutral instructions also were studied. These Ss were told that List B would be a practice-list, and that their performance was not the main concern of E, but rather E was interested in studying certain characteristics of nonsense-syllables.⁸

Following the differential instructions, Ss received 20 trials on List B. They were then told to return 24 hr. later for another part of the experiment, when they were given 10 more trials on List B. E was unaware of which set of instructions and which list S was to receive until the 15 trials of List A had been completed.

Results. To test for the possibility of preexperimental differences among the nine groups, a 3×3 analysis of variance was performed on scores for List A. The measure employed was the number of correct anticipations. The results of the analysis showed that the groups did not differ significantly among themselves. To control for individual differences in learning ability, the results for List B were evaluated by means of analysis of covariance, with the number of correct anticipations on List A as the control variable.

An analysis of covariance on the number of correct anticipations on the 20 trials of List B (Day 1) revealed significant *F*s for associative value ($p < 0.001$) and instructions ($p < 0.01$). Table I presents the means and adjusted means for per-

⁶ Sarason, *op. cit.*, *J. exp. Psychol.*, 51, 1956, 253-260.

⁷ *Ibid.*, 253-260.

⁸ *Ibid.*, 253-260.

formance on List B, Day 1. Comparisons among the adjusted means for the three levels of associative value showed that the groups which learned the list of 80% associative value performed at a significantly higher level than did the groups which learned the 53% and 27% lists. The latter two lists did not differ significantly from each other. Comparisons among the adjusted means for the three groups differing in the kind of instructions administered showed that the groups which learned under the neutral control condition performed at a significantly lower level than did the subject-oriented and experimenter-oriented groups. The latter two groups did not differ significantly from each other. The interaction between associative value and differential instructions was not significant. An analysis of the results for Trial 20

TABLE I
MEANS AND ADJUSTED MEANS FOR NUMBER OF CORRECT
ANTICIPATIONS, TRIALS 1-20, LIST B

Associa- tive value (%)	Neutral		Subject-oriented		Experimenter-oriented	
	mean	adjusted mean	mean	adjusted mean	mean	adjusted mean
27	91.1	88.8	112.7	104.1	88.1	99.7
53	87.9	85.6	115.7	118.0	113.6	112.8
80	156.1	145.8	166.3	178.5	162.7	161.0

of Day 1 yielded essentially the same results as those obtained for the analysis of Trials 1-20 on Day 1. The trends of the learning curves of the three associative-value groups on Day-1 were not statistically different. A similar lack of significance was obtained when the three motivational groups were compared.

An analysis of covariance was also performed on the number of correct anticipations on Day 2 (List B) to evaluate the effects of the experimental treatments 24 hr. after their administration. The only significant *F* obtained was that for associative value. Again groups learning the list of 80% associative value were superior to the groups which learned the 53% and 27% lists.

Another analysis of covariance of performance on Day 2, using the number of correct anticipations on List B (Day 1) as the control variable was also employed. Again, the only significant effect was that for associative value, but comparisons among these adjusted means showed the groups learning the 27% and 80% lists to be superior to the group which learned the list of 53% associative value.

Discussion. It seems likely from the present results that associative value is not linearly related to performance in verbal learning. The fact that no significant difference was found between the 53% and 27% lists does not invalidate Glaze's lists, but suggests the need for further study. The present results concerning the effects of the differential motivating instructions confirm previous findings that subject-oriented and experimenter-oriented instructions have a facilitative effect on the performance of moderately-anxious Ss.⁹ In both previous experiments the subject-oriented instructions tended to lead to a higher level of performance than did the experimenter-oriented instructions, and it is indeed quite surprising to find that

⁹ *Ibid.*, 253-260; *op. cit.*, *J. abnorm. soc. Psychol.*, 166-171.

merely requesting Ss to assist *E* by trying hard was almost as effective in the present experiment as telling Ss that they were taking an intelligence test.

Two results are deserving of further comment. One finding leads one to wonder whether control of anxiety alone is sufficient to reduce the inter-S variability so often observed in studies of experimentally manipulated motivation. When scores on the Taylor Anxiety Scale were manipulated in a previous experiment, differential motivating instructions and failure reports led to no increase in inter-S variability.¹⁰ In the present experiment, although Bartlett's test did not lead to rejection of the null hypothesis, there were considerably more large differences in variance among the nine groups of Ss.

The second finding of interest concerns the absence of a significant effect for instructions on Day 2, although previous experiments have shown a significant effect in delayed tests of retention.¹¹ The differences in the adjusted means were, however, in the same direction as those in the earlier studies. Because the interaction of associative value and instructions was not significant, *ts* were not calculated between pairs of groups, but it appeared that the motivating instructions tended to lead to superior performance in the groups which learned the 27% and 53% lists. No trend in this direction was observed in the 80% groups. Further experiments are indicated.

Summary. This experiment investigated the effects of associative value and differential motivating instructions on verbal learning. Level of anxiety was controlled by selecting the Ss from those individuals falling into the middle range of the Taylor scale. The Ss who learned a list of nonsense-syllables of 80% associative value were superior in performance to Ss who learned 53%-27% lists. Both experimenter-oriented and subject-oriented instructions led to performance which was superior to that of a neutral control group. When retention was measured 24 hr. later, only the variable of associative value was found to be significant. The interaction between associative value and instruction was insignificant.

¹⁰ Sarason, *op. cit.*, *J. exp. Psychol.*, 253-260.

¹¹ *Ibid.*, 253-260; I. G. Sarason and B. R. Sarason, The effect of motivating instructions and reports of failure on verbal learning, this JOURNAL, 70, 1957, 92-96.

MEASUREMENT OF CUTANEOUS ELECTRICAL THRESHOLDS IN ANIMAL RESEARCH

By ALFRED B. SHAKLEE, University of Denver

Though electrical stimulation has been widely used as a motivating agent in animal work, liminal intensive values for a given species have been rarely established. The earliest work in this area is apparently Muenzinger and Mize's study on rats.¹ Using ascending and descending series of intensities, and stimulating the animal only when all four of its feet were touching the electrical grid, these authors obtained a mean threshold for 45 rats (1-6 mo. old) of 0.077 ma. The present study was undertaken to compare two rapid procedures for threshold-measurement.

Apparatus. The box in which the animals were tested was 12 in. long, 3¼ in. wide, and 6 in. high. The grid floor was made of ⅛-in. copper rods spaced at 5/16-in. intervals. A glass plate served as the top. The power source, delivering 60-cycle rectangular, D.C. pulses, produced current whose magnitude was quite independent of circuit-resistance.

Two methods, 'continuous' and 'serial,' were used in applying current to the grid. In the 'continuous' method, the magnitude of the stimulus-current was increased for all Ss at identical rates by means of a constant-speed motor supported on sponge rubber. Control for vibratory stimuli in this type of work is important because of the demonstrated interaction between electrical and mechanical stimulation.² The times required to progress from zero to specified current values were: 0.1 ma., 8 sec.; 0.2 ma., 14 sec.; 0.3 ma., 18 sec.; 0.4 ma., 24 sec.; and 0.5 ma., 30 sec. In the 'serial' method, pulses of 0.05 sec. duration were applied to the grid.

Subjects. Threshold determinations were made on 64 rats (30 males and 34 females) varying in age from 93 to 161 days. Of these Ss, 28 were Wistar albinos and 36 were Wistar-hood crosses. They were randomly assigned to conditions with the restriction of proportional representation in terms of litter and sex. One animal, a male in the 'serial' group, died during the study.

Procedure. After taming of the Ss by preexperimental handling, thresholds were determined on each of three successive days. 'Continuous' and 'serial' groups were divided into subgroups approximately matched in terms of sex and type. It was thus possible on each test-day to use the same counterbalanced running sequence of continuous-serial-serial-continuous. After S was placed in the test-box, 2 min. elapsed prior to the onset of stimulation. Insofar as possible, and with caution not to introduce systematic bias between groups, stimulation was begun when S was relatively quiet. The 'serial' method has advantages in this respect, for in the continuous

* Received for publication February 26, 1957. This study was supported by a research grant (M1527) from the National Institute of Mental Health, and by a University of Denver Faculty Research Grant. The author's thanks are due Mr. David H. Long and Mr. William B. Arndt for technical assistance.

¹ K. F. Muenzinger and R. H. Mize, The sensitivity of the white rat to electric shock: Threshold and skin resistance, *J. comp. Psychol.*, 15, 1933, 139-148.

² J. A. Vernon, Cutaneous interaction resulting from simultaneous electrical and vibratory stimulation, *J. exp. Psychol.*, 45, 1953, 283-287.

procedure the current increase may be started while *S* is quiet only to have him become active while the increase is in progress.

Since it was desired to develop a technique in which *Ss* could be tested within predictable time-intervals, stimulation was not withheld until *S* had all four feet on the grid, as had been done by Muenzinger and Mize. *S* was considered to have responded to the shock when there was a brisk movement of a limb, the trunk, or the head. Changes in breathing-rate could, in many cases, be detected sooner, but because of their ambiguity and unreliability were not considered.

In the 'continuous' method, the motor was stopped upon the appearance of the first clear response. The procedure was repeated over five ascending stimulations and the median result taken as the threshold. There were 30-sec. rest-intervals between determinations.

In the 'serial' method, the first stimulus-value used was sub-liminal, *i.e.* 0.02 ma. Successive increments of 0.02 ma. were given at intervals of 5-10 sec. until *S* responded. When a stimulus-value was reached that yielded a response, stimulation was continued at 30-sec. intervals. Stimulation of this magnitude was repeated until *S* responded on three out of four trials or failed to respond twice. After the criterion of three out of four was achieved in the ascending series, a descending series was begun. At each successive 0.02-ma. step, stimulation was applied until *S* failed to respond on two out of three trials. This was considered a failure, and a value 0.01 ma. higher than the last setting was used. This point was considered the threshold if there were two out of three responses. If this criterion was not reached, the threshold was another 0.01 ma. higher—the last value where there had been two or more responses from three stimulations. The upper criterion was made stringent to insure that the animal was really responding to the current, and the lower criterion was made to hinge on more than one response in the hope of increasing reliability.

A much simpler 'serial' procedure, apparently yielding similar results, has since been adopted in some of our work. The method utilizes an ascending series of 0.03-ma. increments, with approximately 2-3 sec. between stimulations, until a response is obtained. After this, a magnitude 0.01 ma. smaller is applied, and if this responded to, a final value, again 0.01 ma. less, is used. The smallest magnitude responded to is recorded. Five such series are completed, and the median value is taken as the threshold.

Reliability. Coefficients of correlation were obtained by the product-moment method, for the 'continuous' and 'serial' procedures, between threshold scores on Test-Days 1 and 2 and 2 and 3. The results are presented in Table I. The 'continuous' group yielded some significant, though low, reliability coefficients—one for Days 1 vs. 2, and two for Days 2 vs. 3. Using the *t*-test for correlations from the same population, one significant difference was found between the first two-day and last two-day results. This was for males with the pulse-current, the difference being significant at the 2% level.

Threshold magnitudes. Mean thresholds and variances for the different test-days are summarized in Table II. Eight *t*-tests were computed for comparisons involving the first and last test-days. The threshold decrease for all *Ss* from first to last day was significant beyond the 1% level for the 'continuous' method and at the 5% level for the 'serial' method. Comparing threshold-magnitudes obtained by the two

procedures, the difference on the first day was not significant, but on the last day, the 'continuous' procedure gave slightly lower thresholds at the 5% level.

Threshold-comparisons between Wistar albinos and Wistar-hoods for the two techniques separately on first and third days did not yield significant mean-differences. A similar set of four comparisons for males and females, showed numerically lower thresholds for the females in all instances, but in only one case was the difference significant. On the third day with the 'continuous' technique (conditions suggested by the previous data to be more reliable), the mean difference was significant at the 2% level.

Discussion. It is not surprising that threshold-values were considerably higher than those obtained by Muenzinger and Mize. Since fairly rapid techniques were desired, determinations were made when Ss seemed to be making good contact with their hind feet, regardless of whether the forefeet were on the grid. For a large

TABLE I
RELIABILITY-COEFFICIENTS FOR THE TWO METHODS OF THRESHOLD-TESTING

Days	Group	Correlation	
		continuous	serial
1 vs. 2	All	0.36*	0.00
	Males	0.40	-0.27
	Females	0.20	0.30
2 vs. 3	All	0.57†	0.35
	Males	0.72†	0.48
	Females	0.19	0.23

* Significant at the 5% level.

† Significant at the 1% level.

proportion of Ss, if one waited for them to be resting with four feet on the grid for each stimulation, testing-time would be increased many fold. Adaptation-effects of the continuously increasing current, and the multiple-response criteria of the 'serial' method, could both act to increase threshold-values over what other procedures might yield. In interpreting the recorded values, it should be borne in mind that these were based on calibrations with the grid shorted, and not with an S's resistance in the circuit. This would tend to make the obtained scores slightly higher than currents actually being delivered to the Ss. Good stabilizing characteristics of the stimulator would, however, help to minimize such discrepancies. Differences of this sort may, moreover, not be of concern in research where one is interested in determining whether group-means differ, rather than in establishing actual current flow. Aside from the significant differences reported, the consistent numerical decrements in means and variances from Days 1 through 3 seem instructive (Table II).

Thresholds obtained by the methods described are clearly not sensory thresholds. This is made especially obvious when one notes differences between animals in type of response made to small electrical currents. We can, however, consider these measures as varieties of response-thresholds. The observed decrement between first and third test-days seems worthy of further study. It has at least an interesting superficial parallel to known practice effects on reaction-time and on sensory

thresholds.³ That electrical threshold-measures may be useful in differentiating the effects of some experimental variables is seen in a study by Jones where a manually controlled increasing current was used for this purpose.⁴

Conditions which in the present experiment seemed to favor reliability in measuring cutaneous electrical thresholds in the rat were the use of (1) continuously

TABLE II
THRESHOLD MEANS AND VARIANCES (Ma.)

Day	Group	Continuous		Serial	
		Mean	Variance	Mean	Variance
1	All	0.207	0.00393	0.216	0.00525
	Male	0.217	0.00524	0.242	0.00520
	Female	0.198	0.00267	0.194	0.00450
2	All	0.169	0.00426	0.196	0.00330
	Male	0.190	0.00624	0.215	0.00353
	Female	0.150	0.00197	0.181	0.00274
3	All	0.162	0.00187	0.187	0.00250
	Male	0.180	0.00183	0.190	0.00263
	Female	0.146	0.00145	0.185	0.00254

increasing current, (2) male animals, and (3) pretesting. Either of the techniques used with rats in this research, could, with slight modification, be used with other laboratory animals.

³ R. S. Woodworth, *Experimental Psychology*, 1938, 308, 331-332, 358, 454, 483, 506.

⁴ B. E. Jones, Effects of trauma induced stress on learned behavior. Doctoral dissertation, University of Denver, 1955, 17-19.

CALIBRATED MATERIALS FOR VERBAL LEARNING WITH MIDDLE- AND OLD-AGED SUBJECTS

EUGENE H. ROCKLYN, ROBERT B. HESSERT, and HARRY W. BRAUN,
University of Pittsburgh

The authors of this study recently undertook a series of experiments to determine how retroactive inhibition in verbal learning is affected by chronological age. Because most of the studies in this field have been made on college students and with verbal materials that were calibrated and selected upon the basis of their suitability to this age-group (18-23 yr.), we believed that a preliminary study should be undertaken to determine the suitability of the materials to Ss of different ages. Conceivably the responses of Ss of college age are different from those of middle- and old-age.

The variables we examined were *synonymity* and *meaningfulness* of the verbal materials used by Haagen¹ and Noble,² respectively.

SYNONYMY

Haagen classified, upon the basis of meaning, a large number of two-syllable adjectives into 80 groups of 6 words each. From each of these groups one word was selected by 35 judges as the best representative of the meaning of the group. This word was then used as the standard with which each of the other 5 words in the group was paired. A group of 80 Ss (students from the introductory course in psychology) rated the word-pairs for similarity of meaning on a seven-point scale. To determine the reliability of scaling on this dimension, scale-values were computed from the ratings of two randomly chosen groups of 40 Ss. The product-moment correlation computed by Pearson's product-moment method, was 0.92, fiducial limits (1%) 0.88 and 0.93. An example of Haagen's median scale-values for the standard word *wicked* is: corrupt, 2.4; depraved, 3.8; evil, 1.2; sinful, 1.6; vicious, 3.0.

Procedure. From Haagen's groups of adjectives we selected 20 and presented them by the method of paired comparison to 40 Ss—10 from each age-decade between 30 and 70 yr. The Ss in each of these age-groups were divided equally between 'high' and 'low' educational levels, *i.e.* five in each group possessed a college degree and five had never attended college. The education of the 'low' group, all of whom were unemployed and were compensated for their services, ranged between Grades 8 and 12.³ As the word-pairs (standard and one of the five synonyms) were presented, the Ss indicated on a seven-point scale the degree of synonymity of each pair. Median scale-values were computed from their reports.

* Received for publication July 18, 1956. This investigation was supported in part by a research grant, B-787, from the National Institute of Neurological Diseases and Blindness, of the National Institutes of Health, Public Health Service.

¹ C. H. Haagen, Synonymity, vividness, familiarity, and association value ratings of 400 pairs of common adjectives, *J. Psychol.*, 27, 1949, 453-463.

² C. E. Noble, An analysis of meaning, *Psychol. Rev.*, 59, 1952, 421-430.

³ These Ss were obtained through the Pittsburgh Office of the Pennsylvania State Employment Service.

Results. The product-moment correlation between our two groups, *i.e.* high and low educational groups, was 0.79; and between our two groups and Haagen's values, the correlations were 0.79 and 0.77, respectively.

Since the correlation-coefficients of the high and low educational groups were not significantly different, the two groups were combined. The mean age of the combined low and high education groups was 50 yr. Haagen's population consisted of undergraduates, hence their mean age might be reasonably estimated at about 20 yr. A product-moment correlation of 0.82 was obtained between Haagen's scale-values and those secured from the combined high and low educational groups ($N\ 40$). The split-half reliability of the scale derived from the 40 Ss was 0.85.

These results indicate a high degree of relationship between ratings of similarity of meaning made by a group of undergraduate students and similar ratings made by Ss differing in both chronological age and level of education. The correlations we obtained are sufficiently high to warrant the use of Haagen's adjectives with Ss similar to those employed in this study.

MEANINGFULNESS

To obtain a scale of meaningfulness (*m*), Noble presented a list of 96 dissyllables, consisting of nouns and paralogues, to 119 recruits in the U.S.A.F. and 28 undergraduate students.⁴ Meaningfulness was defined in terms of the mean frequency that written associations were made in 60 sec. by the Ss to the stimulus-word. The resulting scale possesses a reliability of 0.975.

Procedure. From Noble's materials, 24 words were selected by taking every fourth one in order of meaningfulness. These materials were presented to 48 Ss—16 in each of the following chronological age-groups: 20-29, 30-49, and 50-66 yr. Level of education ranged from Grade 8 to college graduate with a mean of Grade 12. The Ss, who were unemployed, were compensated for their services.

Continued written associations in a 60-sec. interval were obtained for each of the 24 terms. Responses were scored as acceptable or unacceptable by three judges. Our criteria of unacceptability varied slightly from those used by Noble. His criteria were: illegible responses, perseverative responses, and failures of set. The judges in this study were unable to apply consistently the third criterion and consequently substituted for it, as a criterion of response-unacceptability, S's use of personal names. All other responses were considered acceptable and agreement among the judges was practically perfect. The 24 terms were ordered in mean frequency of written associations.

Results. The mean number of response per stimulus-word made by the 48 Ss was 5.0 as opposed to Noble's mean of 3.5. This difference, which is significant at the 5% level of confidence, may be accounted for by the difference in criteria for response-acceptability described above.

Despite this difference in the mean responses, the product-moment correlation coefficient between Noble's *m*-values and those obtained from the Ss in this study was 0.96. Correlation coefficients between Noble's *m*-values and those of the three

⁴ Noble, *op. cit.*, 426.

age-groups were: 0.96 (20-29 yr.), 0.94 (30-49 yr.), and 0.92 (50-66 yr.), indicating in this case, a tendency for number of responses to decrease with age.

These results indicate that values for meaningfulness of certain verbal materials found by Noble to exist among young adults of college age also prevail among adults who differ in age and level of education.

Summary. This study showed that a high correlation existed between values of synonymity and meaningfulness of verbal learning materials obtained by Haagen and Noble from young adults of college age and comparable values secured from Ss who differed in age and in level of education.

PURSUIT-TRACKING WITH PARTIAL CONTROL OF THE INPUT

By E. C. POULTON, Cambridge, England

In pursuit-tracking the observer has to match an arbitrary input. If the input is repetitive and not too complex, he may be able to anticipate its forthcoming changes, and thus eliminate the normal lag caused by his reaction-time,¹ but he has still, as it were, to follow instructions imposed upon him from outside. In contrast, in one condition of the present experiment, the observer could exert partial control over the input. He was able, as it were, to modify the instructions he received. In this situation, anticipation can assume a new and more active role, and after practice tracking performance should benefit correspondingly. This is the actual state of affairs when a man has to control his vehicle or aircraft, in addition to aiming a missile which he carries at a moving target.

Method. A diagram of the apparatus has been published elsewhere.² An input pointer moved in a vertical dimension over a range of 2.5 in. The movement was the resultant of two simple-harmonic displacements, one twice the size of the other. The subject (*S*) had to use a hand wheel to keep a second (response) pointer in line with the input pointer. The wheel was mounted vertically opposite his right hand in the frontal parallel plane. It had a diameter of 6 in., and needed a tangential force of about 8 oz. to overcome stiction and friction. A rotation of 50° moved the response pointer about 2.5 in.

A second handwheel controlled the frequency of the smaller of the two displacements of the input. It was mounted in similar fashion opposite *S*'s left hand. The wheel was 2 in. in diameter, and needed a tangential force of about 16 oz. to move it. One revolution altered the frequency by 3 c.p.m.

The larger of the two displacements of the input accelerated or decelerated at 6.2 c.p.m./m. In two acceleration trials it increased from 15 to 60 c.p.m., and in two deceleration trials it decreased from 60 to 15 c.p.m. In all trials the frequency of the smaller displacement started equal to that of the larger. A trial lasted 7.25 min.

One acceleration and one deceleration trial were performed under each of two task conditions, according to a Latin-square design. In the single-task condition *S* had to track without any control of the input. The input started as simple harmonic, but became more irregular as the difference in frequency between the two displacements increased. In the double-task condition, in addition to tracking, *S* had to keep the frequency of the smaller displacement about the same as that of the larger displacement, so that the input remained approximately simple harmonic. To achieve this, he was instructed to give the left handwheel three quick turns in the specified direction approximately every 30 sec.

* Received for publication November 12, 1956. This experiment was carried out at the suggestion of Dr. J. A. Leonard. The *Ss* were supplied by the Royal Navy and Royal Air Force. Financial support from the British Medical Research Council is also gratefully acknowledged.

¹ E. C. Poulton, Perceptual anticipation in tracking with two-pointer and one-pointer displays, *Brit. J. Psychol.*, 43, 1952, 226.

² *Op. cit.*, 222.

On one day before the experiment each *S* received about 20 min. of practice on pursuit tracking by itself. He tracked both a simple-harmonic and a more irregular input, covering the range of frequencies used in the experiment. On another day he was taught how to keep the frequency of the smaller displacement of the input about the same as the frequency of the larger displacement. He then received two practice trials in which he had to do this without tracking. One trial involved acceleration, the other deceleration.

Five of the eight *Ss* were enlisted men in the British Royal Navy. The other three were aircrew in the British Royal Air Force. Their ages ranged from 19 to 24 yr.

A continuous ink record of the input and the error was obtained on a paper tape. The first 15-sec. period of each trial was not scored. A score representing the mean error was computed for each minute of the remaining 7 min. of the trial. Units

TABLE I
MEAN ERROR IN TRACKING UNDER THE DIFFERENT EXPERIMENTAL CONDITIONS

Condition	Minute No. 1		Minute No. 7		Minute Ncs. 1 through 7
	approx. input frequency	mean error*	approx. input frequency	mean error†	mean error‡
Single task, acceleration	15 c.p.m.	25	60 and 15 c.p.m.	62§	46
Double task, acceleration	15 c.p.m.	28	? 60 c.p.m.	50§	41
Single task, deceleration	60 c.p.m.	42	15 and 60 c.p.m.	97§	71
Double task, deceleration	60 c.p.m.	45	? 15 c.p.m.	40¶	43

* For both single and double tasks, acceleration different from deceleration at the 1% level of confidence.

† All differences significant at the 5% level or better.

‡ All differences significant at the 5% level or better, except the difference between double task acceleration and deceleration ($P > 0.05$).

§ Minute No. 1—Minute No. 7 $P = 0.01$.

¶ Double task acceleration Minute No. 1 different from double task deceleration Minute No. 7 at the 5% level.

were chosen such that an error of 100 would have been incurred on average without responding. An analysis of variance was carried out, but one-tailed sign tests were used for the differences in Table I.

Results. The results are given in Table I. The order of performing the conditions was not found to be a significant source of variance. The double task of pursuit tracking with partial control of the input produced significantly more accurate tracking than the single task of pursuit tracking alone, on both the acceleration and deceleration trials. The double task had the overwhelming advantage that the input remained more or less simple harmonic all the time. This advantage outweighed the disadvantage that *S* had a second task to perform. Two pilot experiments failed, however, to show this effect because *S* had not received adequate practice on the second task.

With the double task, the significant difference in mean error between the last minute of the deceleration trial and the first minute of the acceleration trial largely

reflects the extent to which *S* could not maintain an approximately simple-harmonic input as the deceleration trial proceeded. It was difficult to maintain a simple-harmonic input, because the display only showed the resultant of the two displacements which were combined in the input. It was possible to tell from the display when the two displacements had slightly different frequencies, but it was not possible to tell which displacement was the faster without some trial-and-error adjustment of the frequency of the smaller displacement. A similar problem is met in compensatory tracking.³

³ *Op. cit.*, 228.

AUDITORY FAMILIARITY AS A DETERMINANT OF VISUAL THRESHOLD

By D. W. FORREST, University of London

In a recent series of experiments, Postman and Rosenzweig investigated the effects of previous auditory or visual familiarity on the visual and auditory thresholds for nonsense-material.¹ They found that frequency of prior training significantly influenced the cognitive threshold, an effect most marked when the prior practice was in the same modality as the later threshold-measurement. The visual test following the auditory training was an exception, however, for there the differences in threshold as a function of frequency were small and insignificant.

The present experiment, which was performed prior to the publication of Postman and Rosenzweig's paper, was limited to a study of the effects of auditory experience on visual threshold and appears to show that auditory familiarity is an important determinant of visual threshold over a wide range of frequencies of presentation.

Procedure. Fifteen Ss (undergraduate women at the University of London) were employed in this experiment, divided for convenience into two groups of seven and eight. They were asked to help *E* in an investigation into methods of learning languages, an artificial language being employed. This 'language' consisted of 18 paralog, or disyllabic six-letter nonsense-syllables, formed by putting together any two nonsense-syllables from Hilgard's list of syllables of low association-value.² In this manner it was hoped to obtain a list of pronounceable words equally difficult to remember. These words were played on a tape-recorder once a week to each group of students. The Ss were told they were to listen to the words and to learn them as thoroughly as possible, and that in a later session they would be given the meanings of these words and tested for retention.³ It was particularly impressed upon the Ss that they were not to write down or otherwise rehearse the words between sessions.

There were five weekly sessions during which the nonsense-syllables were repeated a varying number of times. Six of the syllables were given once a session (making 5 repetitions in all), 6 others were given 5 times a session (making 25 repetitions in all), and the remaining 6 were given 40 times a session (making 200 repetitions in all). Thus, at each session the Ss heard 276 paralog. The time taken to read the complete list was 10-15 min. At each session the order of the paralog was varied.

During the first and third sessions, as a word was spoken for the first time on the tape, a card with that word printed on it was held up in front of the group for 3 sec. Each word thus appeared twice altogether in this written form. This procedure was adopted to familiarize the Ss with the spelling of the words. Apart from these two presentations, the Ss were allowed neither to see the words nor to write them.

* Received for publication November 19, 1956. From the Psychological Laboratory, Bedford College.

¹ Leo Postman and M. R. Rosenzweig, Practice and transfer in the visual and auditory recognition of verbal stimuli, this JOURNAL, 69, 1956, 209-226.

² E. R. Hilgard, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 545.

³ This did in fact occur at a later session, but the results do not concern us here.

The sixth session was used to obtain visual thresholds for the 18 words interspersed with 6 'new' paralogues. The words were printed in India-ink capitals on white cards. A projection-tachistoscope, consisting of an Aldis projector and a Day shutter, was employed. A constant shutter-speed of 0.1 sec. was maintained throughout the experiment, the focus of the lens being changed by 16 steps. The Ss were told beforehand that they would be shown some words with which they might be familiar interspersed with others strange to them, but none of the words would be meaningful. They were asked to make a guess at the word on *each* presentation and these guesses were recorded.

It was impossible in this group-situation to employ a counterbalanced experimental design in which the frequency of occurrence of each word could be varied from one S to the next. Instead, the frequencies were assigned to the various syllables at random. In the analysis of the results this means that peculiarities in structure are not controlled, and the effects of this variable will be confounded with the effects

TABLE I
VISUAL THRESHOLD (MEAN TRIALS REQUIRED FOR RECOGNITION)
AS A FUNCTION OF AUDITORY FREQUENCY

Frequency	Threshold	Difference	t
200	6.2		
		3.4	11.7
25	9.6		
		1.4	3.8
5	11.0		
		3.4	8.1
0	14.4		

of frequency. It is unlikely, however, that the six words of a particular frequency have structural features in common and we are probably safe in ignoring this source of variance.

Results. Table I shows the mean recognitive-thresholds (mean trials required for recognition) of the words grouped according to their frequency of presentation. All differences between pairs of values were significant beyond the 1% level of confidence (*t*-test).

The pre-recognition guesses show that the more familiar nonsense-words were used more frequently as the subject-matter of these hypotheses. A count of the number of different words used as pre-recognition guesses reveals a total of 25 very familiar words (*i.e.* nonsense-words presented 200 times previously), 11 fairly familiar (25 presentations), 11 less familiar (5 presentations), and 2 unfamiliar (words not presented prior to the threshold-estimation). The guesses for the six new paralogues which were doomed always to be incorrect showed the same pattern.

No meaningful words were used as pre-recognition guesses in accordance with the set to look out for and recognize nonsense-words which was induced by the instructions.

Discussion. This experiment appears to have demonstrated that the number of times we have heard words in the past plays an important role in determining the

ease with which we see them in the present. The effect is not limited to the lower end of the frequency-scale as Solomon and Postman have suggested it might be⁴—the fall in threshold continues over the 200 repetitions.

It is difficult to see why auditory familiarity should have had such a marked effect on visual threshold in this experiment in contradiction to that of Postman and Rosenzweig. The experiments are not comparable in several respects: (1) The visual threshold-measurement involved different tachistoscopic procedures—in their case the brightness of a flash of fixed duration was varied, while in our case the focus of the lens was altered. (2) They worked with frequencies of 0, 1, 2, 5, and 15, whereas we used 0, 5, 25, and 200 repetitions. The only direct comparison is between the 0 and 5 repetitions where we found a significant difference and they did not.

We have not attempted to examine the methods employed by our Ss in learning the nonsense-syllables. They were asked to listen to the recordings, but responded in different ways to these instructions. Some were observed repeating the words over to themselves, others closed their eyes tightly or stared at a blank wall. These behavioral differences presumably reflect differences in preferred imagery; the vocalization being especially characteristic of those with strong motor imagery working under a set to remember. It is possible that these imaginal differences would be reflected in the results obtained in the tachistoscopic situation. The visualizers, who had 'seen' each word as it was spoken, might show the effects of familiarity more clearly than Ss without this form of imagery. Such imaginal differences may cut across differences in modality and complicate dual-modality results unless Ss are selected in terms of their preferred imagery.

The finding that our Ss tended to use as pre-recognition guesses the words they had heard most frequently in the past is in agreement with previous results for meaningful words.⁵ In the present experiment, Ss had been instructed that all the words to be presented tachistoscopically were nonsense-words and they had therefore a limited population of expectations from which to choose. Had their instructions not limited them in this way they would probably have begun by supplying common meaningful words as guesses. From an inspection of the protocols it can be seen that Ss soon guessed the source from which the majority of the words exposed were derived, and then used any paralogs they could remember as pre-recognition guesses. This situation is not then very different from a tachistoscopic situation in which S expects meaningful words. In both cases he will supply as guesses words to which he has been exposed most frequently in the past; such words are presumably more readily available as verbal responses, and are therefore produced under the ambiguous stimulus-conditions. In the present experiment the population of possible responses was a limited one (once the source was guessed), but within these limits the most common words were employed most readily—which perhaps accounts, at least in part, for the main effect.

⁴ R. L. Solomon and Leo Postman, Frequency of usage as a determinant of recognition thresholds for words. *J. exp. Psychol.*, 43, 1952, 195-201.

⁵ E. M. McGinnies, P. B. Comer, and O. L. Lacey, Visual recognition thresholds as a function of word length and word frequency. *J. exp. Psychol.*, 44, 1952, 65-69.

THE EFFECT OF SENSORY DEPRIVATION UPON ROTE LEARNING

By JACK A. VERNON and THOMAS E. MCGILL, Princeton University

Several recent studies have been concerned with the effects of various kinds of confinement and isolation upon the intellectual processes of human Ss.¹ In general it has been reported that these processes suffer a temporary deterioration as a result of such experiences. One study, however, has indicated a facilitation of rote learning during confinement.² The present investigation is an extension of that work.

Learning task. Adjective-lists were presented verbally by a tape recorder.³ There were 9 lists, each containing 15 adjectives. The inter-adjective interval was 2 sec. and the inter-trial interval was 5 sec. The method of anticipation was used with a learning-criterion of one errorless trial.

Subjects. The Ss were young men, graduate students.⁴ Of the 12 confined Ss, 9 were able to undergo 72 hr. of sensory deprivation. These were matched with 9 control Ss on the basis of performance on List 1—the Pre-Deprivation Test.

Procedure. Experimental Ss were given List 1 upon reporting for confinement. They were then confined in a dark, light-proof, soundproof, cubicle-like room ($4 \times 9 \times 8$ ft.) which contained only a bed. They wore gloves to which circular cardboard tubes were attached. These tubes extended from below the elbow to beyond the fingertips. Ear-plugs were used to reduce detection of noise of the Ss' own making. They were instructed to remain as quiet as possible since the intent was to reduce all sensory stimulation to as low a level as possible—hence the use of the somewhat inaccurate term 'sensory deprivation.'

After 12 hrs., the Ss were blindfolded and led out of confinement for breakfast and toilet needs. The same procedure was followed for lunch 6 hr. later. After 6 hr. more of further confinement, or 24 hr. after entering the cubicle, List 2 was presented.

* Received for publication February 8, 1957. This research was conducted under contract DA-49-007-MO-671 between the Office of the Surgeon General of the Army and Princeton University.

¹ H. Azima and F. J. Cramer, Effects of partial isolation in mentally disturbed individuals, *Dis. Ner. Sys.*, 17, 1956, 3-8; W. H. Bexton, W. Heron, and T. H. Scott, Effects of decreased variation in the sensory environment, *Canad. J. Psychol.*, 8, 1954, 70-76; W. Heron, B. K. Doane, and T. H. Scott, Visual disturbances after prolonged perceptual isolation, *Canad. J. Psychol.*, 10, 1956, 13-18; Heron, The pathology of boredom, *Sci. Amer.*, 196, 1957, 52-56; J. C. Lilly, Mental effects of reduction of ordinary levels of physical stimuli on intact, healthy persons, *Psychiat. Res. Rep.*, 5, 1956, 1-3.

² J. A. Vernon and John Hoffman, Effect of sensory deprivation on learning rate in human beings, *Science*, 123, 1956, 1074-1075.

³ The adjectival lists were constructed by using the first word in two-syllable adjective-pairs offered by E. R. Hilgard, in S. S. Stevens, *Handbook of Experimental Psychology*, 1951, 548-552.

⁴ Although the Ss were paid at the rate of \$20 a day, they were somewhat difficult to obtain for the reason that most graduate students find it difficult to give up their work for 4 or 5 days. It is of importance to note that all the control Ss were willing to serve except that they were not able to so interrupt their work schedules.

The Ss were blindfolded and seated on the foot of the bed while the recorder played the list in the antechamber of the confinement-cell. Thereafter, during confinement, adjective-lists were presented at 12-hr. intervals coinciding with breakfast and dinner. There were thus a total of 5 Deprivation-Tests presented, the Ss being released immediately following the last, *i.e.* List 6. After 24 hr., the Ss reported for List 7. Lists 8 and 9 followed at 12-hr. intervals. These three adjective lists comprised the Post-Deprivation Tests.

The control group learned the same lists, in the same order, on the same time schedule, and in the same room as the experimental group.

Results. Table I presents the results in terms of mean differences (experimental minus control). The measure of significance was a *t*-test for paired scores.

When trials to criterion or errors to criterion were considered there was no indication of intellectual deterioration on the part of the experimental group rel-

TABLE I
MEAN DIFFERENCES: EXPERIMENTAL MINUS CONTROL

	Pre-deprivation	Deprivation	Post-deprivation
Trials to criterion	0.4	- 2.4	- 0.7
Error to criterion	-0.3	13.5	10.2
Overt errors (%)	3.85	- 5.70*	- 2.67
Fluctuation-cycles	-3.2	-11.2*	0.0
Efficiency-ratio	0.06	- 0.46†	0.22

*= $P < 0.01$

†= $P < 0.02$

ative to the control. Data on trials to criterion also failed to support the previous finding in this laboratory.⁵ The earlier study utilized only 48 hr. of confinement, however, and when the present data through the 48-hr. period were examined there was a slight, but insignificant difference which favored the experimental group.

An overt error is an incorrect verbal response, as opposed to a covert error which is lack of any verbal response. Approximately 20% of the Control Group's errors are of the overt type in all three groups of tests. The Experimental Group differs from the Control Group only in the Deprivation-Tests, where they made fewer overt errors.

A "fluctuation cycle" was early defined by Hull as "one or more consecutive successes" (on a particular adjective in a particular list) "followed by one or more consecutive failures."⁶ Thus, there may be several fluctuations on each of the 15 adjectives in each list before the learning-criterion is reached. Table I shows significantly fewer cycles for the Experimental Group in the Deprivation-Tests.

A related measure of fluctuation at the threshold of recall (efficiency) may be defined as the difference between the number of trials required to reach criterion and the number of trials required to get every item correct at least once, divided by the number of trials to criterion. For example, if a particular S correctly anticipated all 15 items in a list at least once by the end of the tenth trial, but reached criterion only on the fortieth trial, his efficiency-ratio would be:

⁵ Vernon and Hoffman, *op. cit.*, 1075.

⁶ C. L. Hull, The formation and retention of associations among the insane, this JOURNAL, 28, 1917, 419-435.

$(40 - 10)/40$ or 0.75. Thus, the lower the ratio, the greater the efficiency. By this definition, the Experimental Group was significantly more efficient than the Control Group on the Deprivation-Tests alone.

A further analysis was performed on the data for trials to criterion: an *F*-ratio was used as a test of variability. No significant differences were found between the Pre- and Post-Deprivation Tests. On the Deprivation-Tests, however, the Experimental Group was found to be significantly less variable ($P < 0.05$) than the Control Group.

After reaching criterion on List 9, all the Ss were required to relearn all the lists. The resulting savings yielded no significant differences between the two groups on any of the measures in Table I.

Summary. The effect of 72 hr. of sensory deprivation on the learning of lists of adjectives presented auditorially was investigated. There was no significant difference between experimental and control groups in trials or errors to criterion. On the Deprivation-Tests, the Experimental Group made significantly fewer overt errors and showed less fluctuation at the threshold of recall. The Experimental Group was significantly less variable than the Control Group on the Deprivation-Tests. None of these significant differences persisted, into post-deprivation testing.

PARADOXICAL WARMTH: A SUFFICIENT CONDITION FOR ITS AROUSAL

By WILLIAM LEROY JENKINS, Lehigh University, and A. CHARLES KARR,
The Franklin Institute

Paradoxical warmth, the experience of warmth aroused by temperatures below physiological zero, *i.e.* by cold stimuli, has not attained the recognition accorded its counterpart, paradoxical cold. In his recent text, Geldard notes that few investigators have succeeded in producing paradoxical warmth and suggests that what appear to be positive instances may actually result from the mechanical component of stimulation.¹ Perhaps it was this strong suspicion that led Geldard to omit reference to a study, reported by Pavlicek and J. G. Jenkins in 1933, in which the stimulation of previously mapped warm spots by temperatures of 39° and 26° C yielded reports of paradoxical warmth by 29 of 33 *Os*.² These results bulk too large to be dismissed without serious consideration.

Two objections may, however, be raised against accepting the results of this study as definite evidence for paradoxical warmth. (1) When a stimulator is applied by hand, as Pavlicek and Jenkins' was, distortion of the skin does take place and purely mechanical stimulation may occasionally evoke a report of warmth, as Geldard suggests. (2) Because of the interspersion of normally warm stimulations (39° C) among the normally cold (26° C), some of the *Os* may have acquired a set for reporting warmth and may occasionally have given such a report even when warmth was not actually experienced. Are either of these objections fatal to Pavlicek and Jenkins' conclusions? Would either or both operating together lead to reports of paradoxical warmth by 29 of 33 *Os*, if the phenomenon did not exist?

Unequivocal evidence of paradoxical warmth can be provided by stimulating the same spot with a range of temperatures and obtaining (1) several reports of warmth for a range of temperatures above physiological zero, *i.e.* the neutral temperature; (2) several reports of indifference for a range of temperatures around physiological zero; and (3) several reports of warmth for a range of temperatures below physiological zero.

Neutral reports around physiological zero, with positive reports at stimulus-temperatures both above and below physiological zero, would provide a definite safeguard against the artifacts of mechanical stimulation and of suggestion. It is necessary, however, to obtain reports of paradoxical warmth before it is possible to apply a crucial test of their genuineness. In what began as a study of the limen of warmth, a simple way of obtaining reports of paradoxical warmth was discovered, thus providing opportunity for applications of the crucial test.

Method. The apparatus and procedure were similar to those used by the senior author in analogous studies of cold-sensitivity.³ A circular metal field, thermostatically maintained at physiological zero (skin-neutral), is lowered onto the volar surface of the forearm which is held in an arm-rest. Through a hole in the center of the neutral field, a 3-mm. stimulator is moved down by a cam until its tip is

* Received for publication April 26, 1957. [Professor W. L. Jenkins, an ardent investigator in the field of cutaneous sensitivity, died on July 25, 1957. His loss will be keenly felt by his co-workers.]

¹ F. A. Geldard, *The Human Senses*, 1953, 223.

² George Pavlicek and J. G. Jenkins, Paradoxical warmth, this JOURNAL, 45, 1933, 353.

³ W. L. Jenkins, Adaptation in isolated cold spots, this JOURNAL, 49, 1937, 1-22.

exactly flush with the surface of the field. Since the skin normally is not distorted, *O* reports no impression of contact as such. If the temperature of the stimulator is higher than that of the skin, *O* may report an experience of warmth of Intensity 1, 2, or 3. If he reports 'no warmth,' *i.e.* Intensity 0, the field is moved to a new area of skin and another trial made. When a positive report is obtained, the limen for warmth is determined using as few stimulations as possible to secure reports of Intensity 1 and Intensity 0 with stimulus-temperatures differing only by a degree or so.

In the present study, determination of the normal limen for warmth was followed by a 10-sec. adapting contact of the stimulator at either a higher or a lower temperature, after which the limen was determined once more. When it was discovered that low-temperature, following the original liminal determination, was yielding numerous reports of paradoxical warmth, the procedure was modified as outlined above.

Observers. Sixteen *O*s, unpaid volunteers, naïve regarding this field of research and untrained in the observation of temperature-sensitivity, were used in this study. Because they were giving their services and were pressed for time by other duties, many *O*s were used and only a few experiments were conducted with each of them.

Results. Of 32 warm spots stimulated for 10 sec. with temperatures above physiological zero, 29 showed depression of sensitivity to such an extent that a mean temperature of 4° C above the normal limen was required to yield a minimal (Intensity 1) experience of warmth. Of 33 spots stimulated for 10 sec. with temperatures below physiological zero, 28 were enhanced to such an extent that a mean temperature 3° below the normal limen was now sufficient to yield a minimal experience of warmth. These results are analogous to those previously obtained with cold-sensitivity.⁴

No reports of paradoxical cold were obtained in response to the 10-sec. stimulations with temperatures above physiological zero, nor during any of the exploratory stimulations in the search for warm spots. Eight spots on 6 *O*s yielded paradoxical warmth, however, in response to the 10-sec. stimulations with temperatures below physiological zero.

The crucial test was then applied to 74 spots on 16 *O*s. There were 18 spots on 10 *O*s that met the criteria of several reports of normal warmth, several reports of neutral, and several reports of paradoxical warmth. Particularly striking is the intensity of warmth reported from paradoxical stimulation. Of the 67 paradoxical-warmth responses, 33 were reported at Intensity 3, the highest level.

Why was paradoxical warmth so readily obtained by Pavlicek and Jenkins, whereas other investigators have reported negative results? From our study, it appears that paradoxical warmth occurs *frequently* only when the spot has just previously been subjected to normal warmth-stimulation. In the study of Pavlicek and Jenkins, warm spots were stimulated alternately with temperatures of 39° and 26°C. Their aim was to obtain comparable reports for normal and paradoxical warmth from the same spots. Yet in stimulating first for normal warmth and then for paradoxical warmth, they set the stage for the frequent appearance of a phenomenon that other investigators have found only on rare occasions. Now, a quarter of a century later, we are able to verify and to confirm their results.

⁴ Jenkins, *op. cit.*, 1-22.

NOTES AND DISCUSSIONS

ON RECENT STUDIES OF THE LAW OF EFFECT IN INCIDENTAL LEARNING

Evidence that rewards do not strengthen stimulus-response connections under conditions of incidental learning was presented in a well-known experiment of Wallach and Henle.¹ These investigators used a typical Thorndikian procedure in which *S* was rewarded and punished by announcements of 'right' and 'wrong' for numbers guessed in response to a series of words. Intent to learn was presumably eliminated by informing *S* that the experiment was concerned with extra-sensory perception (*ESP*), and that responses called 'right' on one trial might or might not be correct on the next trial. Under these conditions rewarded responses were neither repeated nor recalled more often than punished ones. The stimulus-list was presented a large number of times, however, and different responses to identical stimuli were reinforced on successive trials. Thus, associative interference may have contributed to the negative results. In a recent repetition of the experiment, Postman and Adams used only one training-trial followed immediately by a test-trial.² Repetition and recall were significantly higher for rewarded than for punished responses. These results threw doubt on the conclusions of Wallach and Henle and indicated that rewards may be effective in the absence of intent to learn.

Since the publication of the Postman-Adams study, two investigators—Bitterman³ and Porter⁴—have taken issue with some features of its procedure and its conclusions. In the interest of clarifying the present status of the problem, their criticisms should be met. Since Porter's paper raises some general methodological issues, it will be considered first.

Porter's study. The main purpose of Porter's study was to evaluate the effects of reward and punishment in an *ESP* situation with two different rates of presentation. Postman and Adams had used a 5-sec. rate, and Wallach and Henle a 2-sec. rate. In reviewing the work of Postman and Adams, Porter listed four factors which he

¹ Hans Wallach and Mary Henle, An experimental analysis of the law of effect, *J. exp. Psychol.*, 28, 1941, 340-349; A further study of the function of reward, *ibid.*, 30, 1942, 147-160.

² Leo Postman and P. A. Adams, Performance variables in the experimental analysis of the law of effect, this JOURNAL, 67, 1954, 612-631.

³ M. E. Bitterman, Information and effect in incidental learning, this JOURNAL, 69, 1956, 410-416.

⁴ L. W. Porter, The effect of "right" in a modified Thorndikian situation, this JOURNAL, 70, 1957, 219-226.

believes "might have spuriously increased repetition and recall of rewarded responses."⁵

(1) In the Postman-Adams study, the same three serial positions—6, 13 and 20—were always rewarded, "thus giving rewarded responses any possible *advantages* associated with those particular serial positions."⁶ The series consisted of 25 items, but only responses to Items 3–23 were actually scored in order to eliminate gross influences of primacy and recency. It would be an unusual curve of serial position, in which the mean number of repetitions in Positions 6, 13, and 20 would be significantly higher than the average for all the other positions.

Fortunately, some empirical evidence on this question is available in a study by Tilton, who presents curves of serial position for responses which have been neither rewarded nor punished.⁷ In Tilton's curve for a series of 25 items, the mean percentage of repetitions of Items 6, 13, and 20 is 42.0; the mean percentage for all other items between Positions 3–23 is 42.5.⁸ The suggestion that the items rewarded by Postman and Adams benefited from their particular serial positions appears to be unfounded. It may be noted in passing that these serial positions were chosen for purposes of controlling and assessing spread of effect. Thus, the distances between successive rewards were constant, and there were at least three punished responses preceding and following each reward.⁹

(2) While the serial positions of the rewarded words were the same on the training- and test-trials, the order of the punished words was changed. Rewarded responses may, therefore, have had the advantage of more distinctive positional cues. This argument has considerable plausibility at first glance but does not stand up under careful analysis. It is true that in the case of rewarded responses the maintenance of constant serial positions may favor repetition. It is also probable, however, that *changes* in serial position favor *repetition* of the punished responses. Announcement of 'wrong' leads to variability of response, *e.g.* a disposition to give a number other than the punished one.¹⁰ Shift in serial position removes one of the cues to response change. Thus, repetition of 'wrong' as well as 'right' responses was favored.

(3) All of the Ss in the Postman-Adams experiment were given recall-tests and later informed of the purpose of the experiment. Porter suggests that "communication to potential future Ss was, therefore, a real possibility."¹¹ This inference is contrary to fact, as it was explicitly stated that none of the Ss knew the purpose of

⁵ Porter, *op. cit.*, 220.

⁶ *Loc. cit.* Italics ours.

⁷ J. W. Tilton, Gradients of effect, *J. genet. Psychol.*, 66, 1945, 5-9.

⁸ Estimated from Tilton's Curve B, *ibid.*, 9. Tilton's data are for syllable-responses rather than number-guesses. Since a multiple-choice arrangement was used, his results probably provide a fair estimate of serial-position effects in the experiments under discussion.

⁹ A gradient of repetition following the rewarded positions was obtained.

¹⁰ According to Thorndike, punishment favors variability of behavior even though such variability may be masked by the effects of exercise. See *Fundamentals of Learning*, 1932, 276 f. The fact that failure induces variability of response has also been stressed by proponents of the contiguity position. See V. W. Voeks, Postremity, recency and frequency as bases for prediction in the maze-situation, *J. exp. Psychol.*, 38, 1948, 495-510.

¹¹ Porter, *op. cit.*, 220. See also Bitterman, *op. cit.*, 415.

the experiment.¹² Since this statement does not appear to have been sufficient, we must now indicate the basis of the assertion: (a) All Ss were questioned at the end of the experiment concerning their expectation of a test of repetition or recall. (b) All Ss were pledged not to discuss the experiment with others. (c) For each of the nine groups in our experiment, separate analyses were made for the first and second half of the sample. If there had been substantial communication among Ss, the performance of the second half should have been systematically higher than that of the first half. Nowhere in the data was there any evidence of such an effect. Both Porter and Bitterman attempted to control communication by omitting tests of recall and not giving Ss any information at the end of the experiment.

(4) Since 18 punished words were compared with 3 rewarded words, it was not possible to use each word equally often as a rewarded and punished item. Two forms of the list were used, however, and each of six critical words was rewarded for one-half of the Ss and was punished for the other half. Porter expresses concern that "these six particular words might by chance have stronger initial word-number associations than the always-punished words."¹³ The Postman-Adams results show no evidence of such a bias nor do Porter's.¹⁴ Nevertheless, the criticism is formally valid. Porter attempted to meet this difficulty by limiting analysis to three blocks of four words each, within which rewarded and punished words as well as the serial positions of the two classes of items were counterbalanced. This is probably an improvement but raises another difficulty. Limitation of the analysis to three blocks of four words each reduces the average distance between the rewarded and punished responses included in the analysis. To the extent that there is a spread of effect which is not entirely accounted for by guessing-sequences, the difference between rewarded and punished responses would thereby be reduced.¹⁵

Porter's results substantially confirm the findings of Postman and Adams, although the difference between rewarded and punished responses was smaller than in their experiment. Rate of presentation was not a significant variable, and for the combined groups the differential effect of reward was significant. In the light of our analysis of Porter's criticisms, we do not consider it likely that the quantitative difference between the two experiments may be ascribed to the elimination of "procedural artifacts" in the earlier experiment. The effect of reward in a Thorndikian situation is complexly determined and variations among comparable experiments are to be expected. Instead, we prefer to emphasize the basic agreement between the results of the two studies.

(5) To evaluate the role played by the perceptual isolation of infrequent 'right' responses embedded in a series of 'wrong' responses, Postman and Adams conducted a second study in which the factor of perceptual isolation was varied systematically.¹⁶ Three ratios of 'right' to 'wrong' responses were used: 3:18, 10:11 and 18:3. Rewarded responses were repeated more frequently than punished ones by all groups,

¹² Postman and Adams, *op. cit.*, 616.

¹³ Porter, *op. cit.*, 220.

¹⁴ Personal communication from Dr. Porter.

¹⁵ For a review of the evidence which suggests that spread of effect may not be entirely accounted for by guessing-sequences see M. H. Marx, Spread of effect: a critical review, *Genet. Psychol. Monogr.*, 53, 1956, 119-186.

¹⁶ Postman and Adams, 'Isolation' and the law of effect, this JOURNAL, 68, 1955, 96-105.

and the differential effectiveness of rewards did not vary significantly from condition to condition. It may be noted that the results for the 3:18 group duplicate the results of the first experiment quite closely. Porter describes this experiment as "not entirely convincing" because (1) the difference in frequency of repetition was relatively small for the 10:11 group, and (2) "no mention was made of how the group with 18 'right' interpreted an ESP situation in which chance was 10% but in which they were 'right' about 85% of the time."¹⁷ Three comments are in order here: (a) The statistical analysis unequivocally supports the contention that there was no significant difference among the groups. In addition, we have discussed in some detail the conditions of intraserial interference which probably account for the relatively small difference in the 10:11 group. Porter makes no reference to that analysis. (b) The critical comparisons are between the 3:18 and 18:3 groups which make it possible to compare rewarded and punished responses with isolation held constant. These comparisons reveal major differences in favor of rewarded responses. (c) The reference to the interpretation of the situation by the Ss is obscure, but the following points are probably relevant. As stated in our report, Ss in the 18:3 group were warned beforehand that long runs of 'rights' might occur by chance.¹⁸ Prior to the end of the training trial, Ss had no information about the length of the series (except perhaps by ESP!) and, therefore, could not know that their final score was to be 85%. In any event, it is not clear why possible suspicions about the correctness of E's replies should eliminate the effects of perceptual isolation. Finally, our post-experimental inquiries show that Ss interpreted the high frequency of 'rights' either as evidence for their success in ESP or as the kind of chance run described in the instructions. On the basis of our results we would continue to insist that perceptual isolation is not a major determinant of the differential effectiveness of rewards.

Bitterman's study. The main purpose of Bitterman's study was to show that information rather than effect accounts for the difference in repetition of 'right' and 'wrong' responses. His experiment comprised three groups: (1) a control group for which no responses were called 'right' or 'wrong'; (2) a group tested according to the procedure of Postman and Adams; and (3) a corrected group in which announcements of 'right' were made as usual, but each announcement of 'wrong' was followed by the specification of the correct response. Both experimental groups showed a tendency to repeat more rewarded than punished responses, but the effects were small and not significant. The reason for the discrepancy between the effects of 'right' in Bitterman's study and ours is not immediately apparent. The most important finding in Bitterman's experiment, however, was the significant tendency for Ss in the corrected group to shift from the responses called 'wrong' to those specified as correct by E. Since correction cannot be considered as rewarding, Bitterman concludes that information *per se* produced significant changes in Ss' responses. Information that a response is right may favor learning even under incidental conditions. Hence, it is not necessary to invoke the law of effect to account for the data of number-guessing experiments.

¹⁷ Porter, *op. cit.*, 224.

¹⁸ Postman and Adams, *op. cit.*, 1955, 98.

While Bitterman's analysis may give a plausible account of the results of the corrected group, a serious difficulty is posed by the discrepancy between the effects of a correction and a simple 'right' under otherwise identical experimental conditions. Why should a simple announcement of 'right,' which gives the same information as a correction, be so much less effective, not only in the corrected group but also in the non-corrected group? The effectiveness of the correction may be embarrassing to a simple drive-reduction hypothesis, but the relative ineffectiveness of 'right' in Bitterman's experiment seems to be equally embarrassing to the information hypothesis. Either hypothesis would probably have to be refined considerably before it could account adequately for all of Bitterman's findings.

The issue between reward and information is not likely to be resolved conclusively in experiments using symbolic verbal reinforcements. Our own interest has been primarily in testing the original hypothesis of Wallach and Henle that after-effects do not significantly influence repetition when *S* is not motivated to learn and repetition is not 'sensible.' None of the data reviewed are consistent with that view unless one wishes to insist *a priori* that intent to learn cannot be eliminated. In that case the hypothesis is incapable of experimental test.

University of California

LEO POSTMAN

PAULINE AUSTIN ADAMS

SUBJECTIVE TONES: TARTINI AND BEAT-TONE PITCHES

Two classes of so-called "subjective tones" exist: Tartini pitches¹ and beat-tone pitches.² I believe that no others exist and that those two have entirely different physiological bases. "The viewpoint of the biologist," as I recently stated, "suggests that there is a mutual dependence between the development of the motor organ of 'signaling,' *i.e.* the human larynx, and the development of the sense organ for perceiving these human signals, *i.e.* the human cochlea."³ "When we combine two physical frequencies in order to listen to Tartini pitches, we get into difficulty if the pitch-span of total perception passes upwards beyond the limits of human-voice frequencies. Evolution may have brought it about that there is a peculiar relation between speaking and hearing to the effect that the human cochlea functions 'quite regularly' only within the frequencies of

¹ Giuseppe Tartini, *Trattato di Musica*, 1753-1754.

² The term 'beat tone' was established as a physiological term by Thomas Young (1773-1829).

³ M. F. Meyer, Observation of a further Tartini pitch resulting from a physical sinusoid added to a Tartini pitch, *J. acoust. Soc. Amer.*, 25, 1953, 1195 b.

human larynxes."⁴ Authorities agree on assigning to all human voices a total span of about three and one-half octaves, extending approximately from 85 to 900 ~. Sopranos who reach 1000 ~ are virtually unknown. This makes it appear rather strange that psychologists experimenting on *typical* auditory reactions nowadays so frequently choose a tone of 1000 ~ as "being centrally located." It goes without saying that *in life* higher pitches, naturally coming from quite different sources, also call for reactions though rarely for typical reactions of the nature of language.

Tartini, using oboes, chose pitches far below 1000 ~ for primary tones; and when Helmholtz checked Tartini's observations with tuning forks, the highest fork he then used, as he reported, was 264 ~. Helmholtz never used any higher forks for *that* purpose.

The usual student in audition is prone, when asked if he has 'learned' anything about 'subjective tones,' promptly to reply: "Yes, they are also called 'combination tones' and are divided by Helmholtz into difference tones and summation tones." This alleged division and this term "combination tones" (which it is now so hard to get rid of) resulted from a *guess* (Vermuthung) by Helmholtz that overloading *the middle ear* could 'mathematically' [sic] result in a difference tone and a summation tone and moreover the whole integrated number series 1, 2, 3, 4, 5, 6, and so forth. Helmholtz nowhere, however, reports that he had heard a summation tone produced by his ear. Neither Tartini with his oboes nor Helmholtz with his forks had a chance to *overload* his middle ear. Acousticians are now in general agreement that the middle ear does not produce subjective tones. We should be consistent and, in order to avoid confusion, discard the statement that "combination tones divide into difference and summation tones." All we need, to speak of subjective tones, are the old terms: 'Tartini pitches' and Thomas Young's slightly more recent 'beat-tone pitches.'

Helmholtz hoped (mistakenly) that the problem of subjective tones might, in the future, be solved by *algebraic equations* (leading to the 'combination tones') like those developed by him in his famous *Beilage XVI* (Appendix 16), where nevertheless he admits "In Wirklichkeit . . . viel complicierter," that is: "The matter is not so simple." Unfortunately, the confused notion of algebraical derivation of subjective tones has been made even more confused by a sort of algebraical derivation invented by Wever and Lawrence. These authors propose that "the ear's distortion is

⁴ Meyer, Observation of the Tartini pitch produced by ' $\sin 9x + \sin 13x$,' *ibid.*, 26, 1954, 561 b.

produced in the cochlea and not in the middle ear. . . . The middle ear mechanism carries out its function of sound transmission with great fidelity."⁵ The latter sentence is in complete agreement with my convictions. To explain the alleged distortion in the cochlea, they propose the following transformation-theory which, they claim, derives from Helmholtz.

A mathematical formulation of the transformation theory is as follows: The displacement varies to some extent with higher powers of P [force or pressure]. Thus $X = a_1 P + a_2 P^2 + a_3 P^3 + a_4 P^4 + \dots + a_n P^n$. Then $\dot{X} = a_1 P_0 \cos \omega t + a_2 P_0 \cos^2 \omega t + a_3 P_0 \cos^3 \omega t + a_4 P_0 \cos^4 \omega t \dots + a_n P_0 \cos^n \omega t$. This derivation gives all the components needed to explain the facts as observed.⁶

Studying the above quotation, I am completely bewildered: first, as to what "the facts" are; secondly, how the last equation can be called a "derivation" in mathematical physics; and thirdly, as to what this has to do with Helmholtz. No such equation, *i.e.* "the sum of the square, the cube, the fourth power and so forth to the n -th power of one and the same cosine," is found anywhere in Helmholtz—and they call this "the transformation-theory of Helmholtz." It is a falsification of Helmholtz.

In reply, I quote from a recent paper, "I recommend, for clearance from Tartini pitches, to undertake observations on beat-tone pitches with primary tones above 1000 ~, and best between 3000 and 5000 ~. . . . I define a beat-tone as *a grouping by the brain* of two or three or more electrical charges received from the sense organ *as if* the group were a unit charge."⁷ This is now systematically applied in the manner of Thomas Young to first rare and then more rapid beat-frequencies, and illustrated by compound curves all the maxima and minima of which had been from common tables *numerically* (not algebraically by formulas) computed. Following the example of Stumpf,⁸ recognizing the existence of two, rather than a single, beat-tone pitches, I call Young's (only) beat-tone "legitimate;" its frequency is the difference directly of the primaries. I call Stumpf's second beat-tone "bastard;" its frequency is the difference of twice the lower primary and the simple upper primary. Both experiment and theory are in agreement and they lead to the following interesting facts: (1) The 'legitimate' one appears as soon as the whirring beats reach about 70 per/sec. The 'bastard' one hesitates until the 'beats' reach

⁵ E. G. Wever and Merle Lawrence, *Physiological Acoustics*, 1954, 172.

⁶ Wever and Lawrence, *op. cit.*, 133-135.

⁷ Meyer, Neurological theory of beat tones, *J. acoust. Soc. Amer.*, 28, 1956, 877-881.

⁸ Carl Stumpf, Bestimmung hoher Schwingungszahlen durch Differenztöne, *Wiedemann's Ann.*, 68, 1899, 107, 109.

hundreds per/sec. (2) They cross, musically speaking, at the Fifth. Stumpf's words, here affirmed, are "Bei der Quinte kreuzen sie sich." (3) The bastard one is inaudible whenever the upper ratio term is an odd number, while the lower term may be odd or even.⁹

The conflict between Tartini pitches and beat-tone pitches can be solved only by assigning all frequencies that are much above 1000 ~ to the sacculus and explaining the facts of audibility through definite and neurologically plausible brain functions. When the primary frequencies are low enough, they penetrate the cochlea hydraulically; and then Tartini pitches are added, while virtually no charges are received by the brain from the sacculus. Let me give a few examples.

Combining the rather high-level frequencies 1680 and 1176 (unit 168) we hear the beat-tone pitches 672 (bastard) and 504 (legitimate). In reduced terms, 10:7 gives 4 and 3. Below 1000 ~ level, however, the frequencies 700 and 490 combined give the Tartini pitch 280. In reduced terms 10:7 gives 4 *but not* 3.

Another example: Combining the higher-level frequencies 1645 and 1175 (unit 235) we hear a beat-tone pitch 470 (legitimate) but no bastard because in 7:5 the higher term is odd. Thus we hear only 2 produced by 5:7. Below 1000 ~ level, however, the frequencies 700 and 500 combined give to us the Tartini pitch 300. That is, in reduced terms, 7:5 gives us Tartini 3, but no such pitch as 2, although 2 is *the difference*. This shows the necessity of ridding ourselves of the totally generalized notion that any and every combination of two frequencies gives two subjective tones, a summation tone and a difference tone.¹⁰ All in all, I am of the opinion that both the sacculus and the cochlea are sense organs of audition.¹¹ In this respect I am a follower of Helmholtz, who was the first

⁹ As long as acousticians stuck to algebraic formulas and despised applying trigonometric-numerical tables, no one was aware of the significance of "odd versus even." Another example of the significance of such combinations is shown in M. F. Meyer, Crucial experiments in cochlear mechanics, this JOURNAL, 66, 1953, 269. The beats there are *doubled only when* one ratio-term is an odd number and the other an even number. Thus the mistuned ratio 1:3 does not display any doubling of the beats.

¹⁰ Asserted in S. S. Stevens and Hallowell Davis, *Hearing*, 1928, 184: "When two loud tones are sounded together, we hear . . . the sums and differences."

¹¹ That leaves untouched the utricle and its semicircular canals as sense organs for straight and rotary motion. It also leaves the utricle as the location of Ménière's disease, a kind of seasickness while on solid ground. Treating it by cutting *partially* the nervus octavus has interesting consequences for hearing of high and low pitches. See, for example, W. D. Neff, The effects of partial section of the auditory nerve, *J. comp. phys. Psychol.*, 40, 1947, 214: "Lesions involving one half or more produced hearing losses for 1000 ~ and above." That is, the utricle and the sacculus are too close, almost inseparately, together.

in physiology to divide tones into those low enough for speech and music and those which are shrill or hissing, and who attributed the latter to the sacculus.¹²

Miami, Florida

MAX F. MEYER

A RATE OF MAKING COMPLEX DECISIONS

Psychological experimentation on choice-responses (decision-making or information-processing as it is sometimes called nowadays) has a long history; some of the earliest studies of this type having been made by Donders around 1868. Within the last few years, however, research in this area has taken a somewhat different turn, largely because of attempts to fit man into the model of a communication system. Within this framework there has been interest lately in the rate at which human beings can make successive choices or decisions. Studies on this problem typically use artificial laboratory tasks of a simple kind—sorting cards, pressing keys, and the like.

Recently I was engaged in a job which in some ways seemed ideally suited for estimating the rate at which a certain complex decision could be made. The task was real, not artificial. It was a task which required special knowledge and for which I was highly motivated. The number of decisions which had to be made was large enough to provide both an adequate period of training and a large number of items for measurement. The mechanics of the job had been simplified for efficiency. Finally, most of the stimulus-objects and the results of all the decisions are open to inspection.

The task was that of classifying about 10,000 references into appropriate categories for a bibliography.¹ Readers who are interested in details concerning the preparation of this bibliography should read its preface. For our purposes we need note only that each reference had been typed on a 3 × 5-in. card. A test-sample of 1200 cards was selected haphazardly from the main collection. These were then distributed into 12 sub-samples of 100 cards each which were thoroughly shuffled before use.

The job proceeded in this way: I read each card and decided in which one of 89 categories it belonged.² This decision was recorded on the card in pencil using a simple letter-number code. The card was then tossed into

¹² H. L. F. Helmholtz, *Tonempfindungen*, 5th ed., 1896, 250.

¹ I. N. McCollom and Alphonse Chapanis, *A Human Engineering Bibliography*, 1956.

² Cross-referencing of the test-items was done after the measurements reported here were finished.

a box for later sorting and filing. Eighty-seven of the 89 categories correspond to the headings in the bibliography.³ The two other categories were for (1) items discarded as not suitable for this bibliography, and (2) items about which I could not decide without further information.

Measurements on the test-sample were begun when the entire job was more than half completed. By that time I had memorized the classification system and had extensive practice in its use. Two new sub-samples of 100 cards each were tested one after the other on each of three consecutive mornings and afternoons. The primary data are the times required to process 100 cards as described above. I tried to work as rapidly as possible without sacrificing accuracy. The rate was a little faster than could be maintained comfortably for long periods of time. Following the test-runs, all of the cards were reexamined for accuracy of classification at a more leisurely pace. Changes were required in 2.2% of the cards suggesting that the test-cards had been processed at about the expected rate. These errors, or changes, were ignored in the calculations which follow.

An analysis of variance showed no significant effects due to (1) days, (2) times of day, or (3) the interaction of days with times of day. For this reason the data for the 12 sub-samples were pooled to yield an average time of 7.35 sec./card. The second of each of the two sub-samples tested on any morning or afternoon was, however, invariably processed faster than the first. The probability that this would happen by chance is $2(1/2)^6$ or 0.03. Thus there appears to have been a genuine speed-up in dealing with the second of each pair of sub-samples. My inclination is to regard this as another example of 'warm-up' which has been widely studied and reported in the psychological literature. In any event, if we regard the six sub-samples which were tested second as being more nearly representative of this job, the data yield an average time of 6.93 sec./card, or an average rate of 0.144 cards/sec. The uncertainty of each choice, computed from the frequency-distribution of 1200 test-items in the 89 categories, is 5.65 bits. Multiplied by 0.144 cards/sec. this yields a rate of about 0.81 bits/sec. In this and in the measurements reported below the possibility of sequential dependencies between cards was ignored.

Chi-squared tests show that the test-sample of 1200 items was not a random sample of the items which appeared in the bibliography.⁴ Com-

³ The bibliography actually has a total of 94 headings, but the entries in Section III-K were not subdivided until just before final editing.

⁴ One of the principal reasons, which is clear to me now, is that the test-sample of 1200 cards had a much smaller percentage of discarded items than did the entire collection of references.

puting an uncertainty for the entire bibliography requires an estimate of the total number of items which were discarded or set aside for further information (because these cards were not actually counted). I have made three such estimates: a 'best guess,' an 'upper limit,' and a 'lower limit' estimate. The three uncertainty values computed for the entire bibliography, multiplied by 0.144 cards/sec., yield a rate of 0.69 bits/sec. with outside limits on this estimate of 0.61 and 0.75 bits/sec.

As one might have expected, the values reported here are very much smaller than the roughly 4 to 7 bits/sec. measured for card sorting,⁵ the 10 bits/sec. measured for key-pressing,⁶ or the 25 bits/sec. measured for talking, typing random texts, or playing random music.⁷

It is important to point out that the measurements reported here have serious limitations, some of which ought to be mentioned. They are for only one person. Measurements such as these are undoubtedly highly personal, depending on the skill, training, temperament, and other characteristics of the subject. Finally, there is no way of comparing the psychological difficulty of this task with other types of decision-making tasks.⁸

The Johns Hopkins University

ALPHONSE CHAPANIS

COMMENT ON THE INFORMATION IN COMPLEX DECISIONS

One point in the results reported by Dr. Chapanis is deserving of further comment. This is that what is being measured can perhaps more properly be called *transmitted* information than anything else, and the system is clearly dealing with more information than is measured as transmitted information.

The problem is that when a person makes a complex decision, as in sorting references for a bibliography, he takes in a good deal more information than is actually utilized in the output. Unfortunately it is quite

⁵ E. R. F. W. Crossman, Entropy and choice time: the effect of frequency unbalance on choice-response, *Quart. J. exp. Psychol.*, 5, 1953, 41-51.

⁶ E. T. Klemmer and P. F. Muller, Jr. The rate of handling information: key pressing responses to light patterns. *USAF Human Factors Operations Res. Labs. Memo Rep. No. 34*, 1953.

⁷ Henry Quastler (Ed.) *Information Theory in Psychology*, 1955, 341-349.

⁸ This study was done in part under Contract Nonr-1268(01), Project NR 145-075 and in part under Contract N5-ori-166, Task Order 1, both with the Office of Naval Research. This is Report No. 166-1-208, Project NR 145-089, under the latter contract.

impossible to estimate how great is this informational input, partly because the input is highly redundant, and partly because the human operator is a curious system who shuts off the input at some point where he decides he has as much information as he wants.

One variation in the procedure that would make this distinction clear would be to restrict the task to a one-bit sorting of the references on some criterion requiring a complex evaluation. Suppose, for example, Dr. Chapanis had sorted the cards into two piles, those primarily of interest to psychologists and those primarily of interest to engineers. I venture to predict that his decisions would not have been made much more rapidly than they were in the present circumstances, and such a series of decisions would have contained roughly one-sixth of the information.

A more abstract version of the same game could be imagined as follows: suppose that you had a list of 500 randomly selected three-digit numbers. Now choose a single number from another set of three-digit numbers and decide whether this particular number is contained among the 500 or not. This would be a one-bit decision, but it would clearly require the processing of ten bits of information. The game could be changed very slightly by requiring the person to sort the target number into one of a thousand cells. He could accomplish this almost as rapidly as he could make the one-bit decision. Yet every item placed correctly in its cell would in this case represent a ten-bit decision.

There is no question at all that many of the decisions made by human operators have the character of Dr. Chapanis' task, that is, they require the transformation of and consequent reduction in the amount of the input information. It is not quite clear to what element in the processing chain a notion of capacity can be meaningfully applied.

E. B. N

AN AUDITORY FIGURE-GROUND EFFECT

In exploring the effects obtained with two alternately sounding tones we found an auditory effect which may be regarded as a figure-ground phenomenon. Under certain conditions, the more intense of the two tones is heard as clearly intermittent (somewhat as 'figure'), and the less intense appears to sound continuously (somewhat as 'ground'). In a sense this is analogous to the situation in vision, where the ground is perceived as extending continuously behind the figure. The purpose of the present note is to call attention to the existence of the auditory phenomenon.

Tones were produced by two oscillators connected through attenuators

to a Grason-Stadler electronic switch and interval-timer. They were then led through an attenuator and a transformer to earphones. The interval-timer controlled the duration and the rate of repetition of the two tones which were sounded alternately.

The figure-ground effect mentioned above appears only when the two tones are brought near together in frequency and when one tone ('figure') is more intense than the other. For example, when repetition is 10 per sec., and one of the tones is set at 4000 ~, with sensation-level of 60 db. and duration of 60 m.sec., and the other tone is set at 45 db. and 40 m.sec. then the second tone is heard to sound continuously when it is approximately 1000 ~ below the 4000 ~ tone; and also when it is approximately 500 ~ above the 4000 tone. The exact region above and below the more intense tone at which the weaker tone is heard as continuous varies with the individual *O*. Shifting the less intense tone as much as 2000 ~ below the 4000 ~ and more intense tone, or as much as 1000 ~ higher than that tone, abolishes the effect and both tones are heard as intermittent. When the more intense tone is at a lower frequency, the frequency regions where the weaker tone is heard as continuous are closer in frequency to the more intense tone than when the more intense tone is at a high frequency.

We intend to establish the quantitative relations for this effect by varying frequency-separation, repetition-rate, duration, and intensity. The results should prove interesting as they may be related to the frequency-limits for simultaneous and successive masking and for the critical bands, in addition to providing another example of the figure-ground effect.

University of Wisconsin

WILLARD THURLOW

THIRTY-SEVENTH ANNUAL MEETING OF THE WESTERN PSYCHOLOGICAL ASSOCIATION

The thirty-seventh annual meeting of the Western Psychological Association was held on the campus of the University of Oregon, Eugene, Oregon, May 9-11, 1957. Two hundred twenty-four persons registered for the convention. The program included 59 papers, 5 symposia, a coffee hour, a business meeting, and the presidential banquet. Business meetings were also held by the Oregon Psychological Association and the Washington State Psychological Association.

The presidential address by Joseph Gengerelli "On the neural events underlying classical and operant conditioning," defended the proposition that the time had come for renewed neurological theorizing in the field of learning and that ultimately learning theory would rest on neuro-

physiological grounds. The Program Committee was chaired by Richard Littman and the Local Arrangements Committee by Norman Sundberg.

At the business meeting the Association accepted an invitation to hold the 1958 meeting at the U.S. Naval Post-Graduate School, Monterey, California, April 24-26th. Dr. Francis Palmer was appointed Program Chairman for 1958. San Diego, California was selected as the site for the 1959 meeting.

Quinn McNemar of Stanford University was elected president-elect for the coming year. The president of the Association for 1958 is Leona E. Tyler.

Occidental College

DAVID L. COLE

ERRATA

Dr. Robert W. Burnham invites attention to an error made in reporting the results of his recent study on "Inter-eye differences in color-perception" (this JOURNAL, 70, 1957, 386-394). In describing the results of observer RE (*Results*, line 10, p. 389), the word "blue" should be 'yellow.' The sentence in question should read, "RE showed significant differences in both studies, and plots of his data show generally that his left eye required more yellow to complete a match."

K.M.D.

Henry Herbert Goddard: 1866-1957

The death on June 18, 1957, of Henry H. Goddard in Santa Barbara, California, at the age of ninety years, probably meant little to many present-day psychologists since he had long been withdrawn both geographically and in spirit from current professional activities. His most important work, moreover, was distant in time and in orientation from the psychological present. In his day, however, he was an important influence not only in psychology but also in sociology and criminology—and in the public mind.

Born in Vassalboro, Maine, August 14, 1866, of Quaker parents, Goddard obtained his A.B. degree from Haverford in 1887 and his A.M. in 1889. After six years as a secondary school principal in Ohio and Maine, he went to Clark University as fellow in psychology in 1896 and received his doctorate there in 1899. From 1899 to 1906 (except for a year of study in Germany in 1903-1904) he was professor of psychology in the Pennsylvania State Teachers College at West Chester. Then came the decisive event in his life—the invitation to be director of research at the Vineland (New Jersey) Training School for the Feeble-minded.

Goddard's twelve years at Vineland may well be called remarkable. There he founded the first laboratory for the psychological study of the feeble-minded and gained for the psychologist recognition of his major value in institutional work. There he translated and adapted the Binet scale and popularized its use in this country; and as a member of the Yerkes committee he helped develop the group-tests of the first world war. There it might be said that, with the fine support of Superintendent E. R. Johnstone, a team approach in psychological research was initiated; the field-worker, the Binet examiner, the house mother, and the psychologist all worked together in appraising cases and carrying through investigations. There the concept of feeble-mindedness was extended to include borderland cases whom Goddard differentiated as morons, and these cases were seen as presenting pervasive socio-economic and educational problems. There extensive data were assembled because they were thought to be indicative of primary feeble-mindedness's being a Mendelian trait. Certain notorious criminal cases were reviewed on the hypothesis that mental defect was the key factor, and appropriate social policy was urged. Educational methods for use with subnormal children were elaborated. Being a friendly out-going man and believing that he had a message of great social significance, Goddard carried his gospel to the public with evangelistic zeal, in lectures and articles and a series of books: *The Kal-*

likak Family in 1912, *Feeble-mindedness* in 1914, *The Criminal Imbecile* and *School Training of Defective Children* in 1915. All were illustrated, full of interesting detail—and widely read.

In 1918 Goddard became Director of the Ohio State Bureau of Juvenile Research, recently set up to study and aid in dealing with juvenile delinquents. While there he published the *Psychology of the Normal and Subnormal* in 1919, *Human Efficiency and Levels of Intelligence* in 1920, and *Juvenile Delinquency* in 1921. Moving to Ohio State University as Professor of Abnormal and Clinical Psychology in 1922, his interests broadened to include a study of multiple personality reported in *Two Souls in One Body?* in 1927, and *School Training of Gifted Children* in the same year. He retired, unwillingly, in 1938.

Brought up in the Society of Friends, he had a fervor which took him into such diverse causes as eugenics and euthanasia and made him sometimes uncritically earnest in his work. Perhaps of all his accomplishments he was most proud of his invention of the word *moron*, feeling that its becoming a part of popular speech meant an enlargement of common understandings—and he was delighted when once (so he declared) he found himself on a Pullman car so named.

He was a fine gentleman, a delightful companion, a great traveler, a seasoned mountain-climber—throughout his career of notably wide-ranging interests and activities. Thus for his first position after graduating from college in 1887 he went into the remote West, to teach in a little place called Los Angeles and a little school ambitiously named the University of Southern California (out near a racetrack which later became Exposition Park); and there he introduced football to the Pacific Coast and coached the first team in that region. In his professional work and in many of his attitudes he was of another generation from the present, but he was an important person in his day, and it might be said that all the social sciences have felt his influence.

The Ohio State University

H. E. BURTT
S. L. PRESSEY

BOOK REVIEWS

Edited by GEORGE L. KREEZER, Washington University

Psychological Research. By BENTON J. UNDERWOOD. New York, Appleton-Century-Crofts, Inc., 1957, Pp. vii, 298.

After five years of teaching a graduate course on scientific method in psychology at Northwestern University, Dr. Underwood presents his lectures in book form. Reflected also are many additional years of the author's experience as a supervisor of graduate students and as a consulting editor for the *Journal of Experimental Psychology*. Underwood presents what he calls his personal philosophy of science. Throughout the book his central objective is the improvement of scientific thinking and research methodology in psychology.

In the introductory chapter, while pointing out that the subject matter applies equally well to both pure and applied research, Underwood strongly defends the need for pure research, about which no one asks the question, "What good is this?" In the second chapter on "The Research Situation," we find the background being laid for later chapters, through discussions of scales of measurement, reliability, and similar subjects. Here, as in later chapters, the author properly places great stress on the reliability of response measures. By contrast, the concept of validity gets only brief and casual treatment, apparently a reflection of the author's primary concern with pure rather than applied research. All of Chapter 3 is devoted to the important subject of "Operational Definitions," a topic which reappears frequently throughout the book.

The main purposes of the book are reached in Chapters 4 and 5, both of which are devoted to "Research Design." The treatment of this area is primarily discussion of types of errors which the author has observed psychologists make most frequently in designing experiments and drawing conclusions from them. Most errors result from confounding among environmental, task, and subject variables, leaving the cause of the observed experimental phenomena indeterminate. Many examples are given of real and fictitious errors in research design.

Most of the remainder of the book is devoted to Underwood's philosophy of science, with particular reference to psychology. Chapter 6 is entitled "An Overview of Explanation in Psychology," and is largely concerned with theory. Opinions of leading psychologists are quoted to show that a difference of opinion exists as to the importance of theory and its proper role in psychology. Chapter 7, "Some Characteristics of Concepts" introduces the author's classification of concepts into five levels along the dimension of abstraction. Level 1 concepts identify independent variables. Identification or naming of phenomena, *i.e.* relationships between dependent and independent variables are level 2 concepts. At the upper extreme of abstraction are Level 5 concepts, which summarize the interaction of postulated processes (e.g. Hull's *effective reaction potential*). In Chapter 8, "Some Explanatory Attempts," the author uses his 5-level classification of concepts to philosophize

about explanation in psychology. A final Chapter 9, aptly named "Potpourri," takes up miscellaneous left-overs.

The overall character of Underwood's book on "Psychological Research" is perhaps easiest to summarize in a negative way by outlining some things which the book is not. First of all, it is not a textbook telling one how to do psychological research, as the title would suggest. That purpose is better served by Underwood's 1949 book on "Experimental Psychology." On the other hand, for persons already fairly sophisticated about psychological research, the book should sharpen their judgment in conceiving, setting up, and describing research. Secondly, the book is not one in which to find help in the use of statistics or mathematics in psychology. Although such quantitative methods are often referred to, the book deals with psychology primarily as a verbal science. Thirdly, I would not call the book a model of clear and concise exposition. In explaining by example, the author too often leaves the reader to search for the principle being illustrated. Nevertheless, it is a valuable book which will deserve serious study, not only by graduate students, but by teachers and advanced research workers as well.

Aero Medical Laboratory
U. S. Air Force

WALTER F. GREYER

Vocational Counseling with the Physically Handicapped. By LLOYD H. LOFQUIST. New York, Appleton-Century-Crofts, Inc., 1957. Pp. XIV, 381.

To the general field of psychology, the area of vocational rehabilitation is a very recent development which, to all intents and purposes, begins with the veterans' program after World War II. Moreover, vocational counseling of the disabled is associated almost completely with the development of counseling psychology. Neither of these premises is exactly true. The literature and practice on counseling the disabled in the State-Federal program for civilians goes back to 1920, and in the field of special education to the 1900's. For various reasons, this early literature has hardly affected the mainstream of psychology. The book by Lofquist may signify a "coming of age" in our present practice of psychology.

The book is the product of experience in the practice of vocational counseling in the setting of a general veterans hospital. With this important fact in mind (for rehabilitation counseling in the setting of the State-Federal program for civilians, of a private rehabilitation center, of special education, or of a public institution, will show real differences in emphasis and scope of practice), the book is scholarly, practical, readable, and capably edited. It portrays the sound application of counseling and psychology for the particular population. The book is a good text for rehabilitation-trainees interested in ultimate employment in a general veterans hospital, but it cannot fulfill the same function for counselors in other rehabilitation settings.

A significant indication of the emerging self-awareness of the rehabilitation counseling psychologist is found in the inclusion of a major section of about 60 pages, or a little less than one-sixth of the book, which deals with the counseling psychologist's relationship to other members of the hospital staff and to outside agencies. This section is a necessary prelude to the book, and to subsequent literature in the same field. The psychologist must find and know his role in the new setting. He must gain acceptance, and convey to other people the nature and scope of his role.

Another major section of the book deals with specific disability areas. The author singles out seven disability groups; diabetic, amputee, heart, hypertensive, paraplegic, cancer and skin. These areas are worthy of review, especially the diabetic, hypertensive, cancer and skin, on the basis of scanty coverage in the literature. On the other hand, the amputee is the subject of many books and articles. The author devotes a chapter to the justification for coverage; this is the weakest chapter in the book. It is claimed that disability groups such as the blind and visually limited, the deaf and aurally limited, the psychiatrically disabled, and the tuberculous, are covered adequately in other literature. The author's major frame of experience in a general veterans' hospital deeply colors the argument, a fact which may be seen in the casual references to other disability areas. Thus, it is said that the available literature on the emotionally disturbed is fairly large. The book deals, however, almost entirely with disease and therapy; very little with the rehabilitation of emotionally handicapped people in the working and social community. The few excellent pieces of literature on the latter topic are not mentioned, such as *Vocational Rehabilitation of Psychiatric Patients*, *Vocational Services for Psychiatric Clinic Patients*, and *Psychiatric Information for the Rehabilitation Counselor*. Nor is there mention of work with the mentally retarded, including the book on *Vocational Rehabilitation of the Mentally Retarded*. Also, the reference to Lende's books about the blind lists the 1940, not the 1953, volume. Reference to the deaf is made in terms of the book by Pintner and Associates published as far back as 1941, although several books have appeared more recently.

The chapter dealing with "The Counseling Process" is characterized by a practical, how-to-do-it flavor. There are many "do's" and "don'ts." It seems to the reviewer that a simple approach to counseling does not stand up under the demands of dealing with people of different disabilities, ages, temperaments and behavior patterns. We still have to develop a general theory of counseling the disabled. The current book portrays the present status of knowledge.

The reviewer does not share the author's conviction that "all counselors should be compulsive about using standardized test directions, timing accurately, and scoring accurately" (p. 283). A more flexible approach is needed in which the tests are regarded as tools to an objective, rather than a slavish end in themselves. In evaluating the cerebral palsied, blind, deaf, emotionally disturbed, mentally retarded, and epileptic, leading psychologists have advocated a "clinical" approach. The experience with projective tests has made it abundantly clear that standardization of testing conditions is not completely attainable. One must consider the psychological and physical state of the individual as well as the specific disabilities that cause interferences with timed testing, conditions which were not contemplated by the constructors of the test in standardizing on a non-handicapped population.

The author states in a few places, and illustrates it in some of the case histories, that attitudes of the family have an important bearing on the counseling of the individual. The function of dealing with family members is left to the social worker in the setting of the general veterans' hospital. In the framework of the State-Federal program of rehabilitation, counselors themselves must deal more actively with family members. Similarly, counselors in the latter program must take a far more active role in job placement of the handicapped than is portrayed in the book.

These critical comments are minor in terms of the many fine qualities of the

book within its real scope of coverage. We mention them here to help the psychologist to see the book in the broader perspective of practice, and to indicate the major limitation of its use in the graduate classroom. The foreword by the distinguished psychologist, D. G. Paterson, states that, "The present volume is the first book to deal systematically with this subject," an extreme statement that can be justified only with many qualifications. The author's Preface is more realistic in presenting the scope of the book.

We believe this book merits wide reading by psychologists interested in the current applications of psychology in work with disabled people. We agree with the author that, "There is much still to be done in the large area of rehabilitation counseling," that a study of the handicapped enriches our knowledge of psychology in general, and that there is a place and need for more books of high quality in this difficult area of scientific application.

Office of Vocational Rehabilitation
U.S. Dept. of Health, Education and Welfare

SALVATORE G. DiMICHAEL

The Nature of Hypnosis. By PAUL SCHILDER. New York, International Universities Press, Inc., 1956. Pp. 204.

Hypnotic Suggestion: Its Role in Psychoneurotic and Psychosomatic Disorders. By S. J. VAN PELT. New York, Philosophical Library, 1956. Pp. 95.

Hypnosis and its Therapeutic Applications. Edited by ROY M. DORCUS. New York, McGraw-Hill, 1956. Pp. x, 309.

A Handbook of Medical Hypnosis: An Introduction for Practitioners and Students. By GORDON AMBROSE and GEORGE NEWBOLD. London, Baillière, Tindall, and Cox, 1956. Pp. xiii, 255.

Interest in hypnosis is cyclical in nature, and the duration of each cycle is said to be some 30 to 50 years. At the moment we are at the peak of such interest, and it is a peak that may continue provided the 'new look' in hypnosis persists. This new look involves greater attention to scientific control, more modest claims for the efficacy of hypnosis, and an emphasis on the adjuvant role of hypnosis in therapy. These four books written by people with diverse interests testify in part to the above facts as well as to the renewed interest in hypnosis present in professional circles.

Schilder's book, republished in the year 1956, was originally written in the years 1920 and 1925 as two separate monographs. While the book is said to be for the layman, the difficulty which the present reviewer found in following some of the points made by the author led him to wonder about himself, or about the translation, or about Schilder. There are many points in the book with which one finds himself in disagreement: the easy equation of sleep and hypnosis, the neurological bases advanced for hypnosis and its varied phenomena, the conclusion that animal and human hypnosis are "basically identical" following a sentence suggesting that the two are necessarily different, and so forth. Schilder also believes that fatigue lowers the ability to resist suggestion, as do occasional intoxications; consequently, he believes that "morphine addicts are as a rule more susceptible to hypnosis during periods of abstinence," an interesting conjecture but only a conjecture, for data are lacking. The presentation is essentially based on psychoanalytic theory, and as such, it encounters many of the objections raised to psychoanalysis

itself; namely, lack of data and reasoning after the facts. In evaluating the book, however, one has to keep in mind that it was actually written some 30 years ago. Although it has been described as being astonishingly up to date, its chief value, in the reviewer's opinion, is historical.

The little book by Van Pelt contains a brief description of many of the phenomena of hypnosis, a brief statement of the author's interpretation of the etiology of the neuroses, and of just how hypnotic suggestions works in effecting a cure. Inasmuch as we are not sure just what hypnosis is, such a description is surprising. One gets the feeling that the author believes that by describing, in diagrammatic form, what is known to exist he has explained the mechanism by which the results obtained have occurred. To describe the functioning of hypnosis in terms of "units of mind" and "doses of suggestion" is not very helpful and a far cry from prediction.

Hypnosis and its therapeutic applications is the theme of the book edited by Dorcus. In this text there are twelve chapters and eight contributors. Of these eight, one is a dentist and the remainder are psychologists. It may be questioned whether in an interdisciplinary topic such as hypnosis, there should be this emphasis on psychology, but the very fact that many of the authors are psychologists undoubtedly makes for more cautious and sounder conclusions because of the greater attention to experimental design. Although the book is described as concerned with therapeutic aspects, some half of its pages are devoted to discussion of such topics as the phenomena of hypnosis, methods of induction, theories, physiological concomitants, genuineness of various hypnotic phenomena, and the like. Although those interested in immediate application may be somewhat disappointed, such discussion is needed inasmuch as the therapeutic value of hypnosis depends in large measure upon our general knowledge of it. This book, while of necessity limited in its overview of the field by the nature of its contributors, belongs on a reading list of the student who wishes to investigate the area of hypnosis. Its many pages conceive of hypnosis not as a panacea but as a useful tool in both diagnosis and therapy.

Ambrose and Newbold have written a book which they claim is not a text but a handbook. It is designed, the authors state, for practitioners and medical students. With the exception of the first three chapters, which are concerned with hypnosis and the law, hypnotic phenomena, and a very brief history of medical hypnosis, the chapters are of the 'show me' variety. The cases described in these chapters are concerned with the use of hypnosis in general medicine, psychopathology, surgery, pediatrics, gynecology, obstetrics, and dermatology. The cases are easy to follow and the book is rich in reference material. While the reviewer is sympathetic with the often repeated statement of the authors that the family doctor is the one best qualified to know and understand the child and consequently to use hypnosis wisely, one wonders how this point of view jibes with the present day increase in and emphasis on specialization. This book is written in the hope that hypnosis may be integrated into the main stream of medicine.

Taken as a group, these four books are interesting in the sense that they show where hypnosis has been, where it is, and where it may go.

State College of Washington

F. L. MARCUSE

The Psychology of Occupations. By ANNE ROE. New York, John Wiley and Sons, 1956. Pp. xii, 340.

A man's occupation is far from independent of his total personality structure. Conversely, his choice of occupation is systematically related to those variables of which personality is a function. Dr. Roe has brought together data which help to clarify and systematize this relation between occupation and personality. She has developed a two-way system of classification in which occupations are arranged into eight groups, depending on the focus of activity, and each group is divided into six levels, primarily on the basis of the degree of responsibility required by the job. The classification, which is essential to the book, is psychological in that it relates fundamental personal characteristics to occupations, something which can not be done with a work-oriented classification such as the one used in the *Dictionary of Occupational Titles*.

The material is organized into five parts: on the role of occupations in society and in the life of the individual; on individual differences, physical, psychological, social, educational, and biographical; on occupational classification and the relation of individual differences to occupational groupings; on following an occupation, with a consideration of the problem of occupational choice, occupational progress, and the effects of occupation on other aspects of living; and on the implications of occupational psychology for education, guidance, therapy, and psychological theory. Relevant literature, representing approximately 500 papers, is surveyed.

One can speculate as to who will profit the most from this book. It was intended "for upper-class college students in programs in vocational guidance, counseling, and clinical psychology." Certainly, high school and collegiate guidance-personnel will find the book valuable, and the students who come for guidance will benefit greatly. Its greatest value, however, may be for personality-theorists who find a new avenue of approach to the study of personality opened up for them.

Knox College

ROBERT S. HARPER

Youth in a Soundless World. A Search for Personality. By EDNA S. LEVINE, New York, New York University Press, 1956, Pp. 217.

The congenitally deaf stand forth as individuals in whom mental, social and even emotional growth are achieved through "artificial" means. Here is provocation for psychologists interested in studies of human development.

Edna Levine's study concerns the results of the administration of the Wechsler-Bellevue Scale for the Measurement of the Intelligence of Adolescents and Adults, Form I, and the Rorschach Method of Personality Analysis of 31 adolescent deaf girls in the Lexington School for the Deaf in New York City. These girls were selected as "normal" from a list of residential girls submitted to three outstanding teachers of the deaf familiar with the girls for the last five years and by the school psychologist. All of the girls had been in attendance at the school for a minimum of five consecutive years. Lexington School is an outstanding residential school employing the Oral Method (speech and lip reading) of instruction. The subjects, who ranged in age from 13 through 18 years, were congenitally deaf or had lost their hearing prior to the age of three, with hearing losses of more than 70 db. in the better ear. Their education had started between the ages of 5-7 years and

their mean reading grade, as measured by the Stanford Achievement Test, was 4.2. As a control group, the norms of 100 children of comparable age, drawn from the original standardization group of the Wechsler-Bellevue Intelligence Test, was used.

In the Rorschach and in the verbal portion of the Wechsler, subjects used a combination of speech, finger spelling and signs, the proportion depending upon the individual.

Results of the Wechsler-Bellevue showed no significant difference in the full scale I.Q. The deaf were inferior in all of the verbal subtests. In terms of the test rationale, the deaf were inferior in knowledge of the world about, in common sense judgment, practical reasoning, general ability to evaluate past experience, in control of attention, and in level of abstract concept-formation. These differences were of degree rather than kind. The deaf were superior on the performance scale, especially in tests of visual-motor coordination and visual organization.

The Rorschach test indicated personality-traits, for the deaf girls, such as emotional underdevelopment, a lag in understanding the dynamics of interpersonal relationships of the world about, highly egocentric life perspective, constructed life area, and rigid adherence to a formal code of etiquette rather than inner sensibilities as standards.

Levine feels that her results give confirmatory evidence of the hypothesis of mental "underdevelopment" and show that personality development is closely linked to language development.

This study of 31 adolescent girls yields results that should stimulate further psychological study. The discussion and concluding chapter indicate that the author's studies and recommendations led to the formation of the Mental Health Center for the Deaf, through the aid of a grant from the Office of Vocational Rehabilitation.

The experimental portion of the study would have made an excellent paper in a psychological journal, and in this form it might well have reached a larger number of psychologists. It becomes a volume through the inclusion of chapters giving surveys of statistics on deafness and methods of teaching, a review of previous psychological studies, especially of the Pintner period, descriptions of the tests used, and illustrative case records.

The study indicates further the difficulty of securing a representative group of deaf subjects and the application of language tests to a group whose chief difficulty is the acquisition and use of language and vocabulary.

Central Institute for the Deaf

HELEN S. LANE

Ageing in Industry. By F. LE GROS CLARK and AGNES C. DUNNE. New York, The Philosophical Library, Inc., 1956. Pp. xi, 146.

A survey of modern industry from the viewpoint of its older operatives is what this particular contribution to ageing is about. What workers are physically able to continue in the various occupations beyond the mid-sixties? What differences appear in the various main occupations of England? To answer questions of this kind, the two investigators have used tables of age and occupations published in the census reports for 1921-1931 and 1951 in Great Britain. The investigators estimate that the population studied by this actuarial approach includes about one-quarter of the occupied male population in 1951. The female population was omitted from the

study. Thirty-two occupations were analyzed. Three sets of what the authors call "exercises" are reported about each occupation: (1) The Moving Cohorts; (2) Age Group Ratio; (3) Change of Age.

The method of "moving cohorts" is given by the authors as the only valid method of determining the number of men who will survive in any occupation to an advanced age. By the study of the statistics in the periods given they can get at least a lump picture of the age group of men at a certain date and discover what happens to all of them through succeeding decades. In tracing the fortunes of what they call a cohort from decade to decade, there are only five possible things described that could have happened to its members in the period that had elapsed since the previous census. They could remain in the same occupation, they could retire, migrate, die or transfer. What the authors do is take the 1921 census figure and move the decennial age groups forward so that they are shown as brought forward under the ages that they would have attained by 1931 had they survived. The numbers actually working in 1931 are then given. By subtracting the lesser from the greater, they then can obtain the measure of the increase or decrease that has taken place in each group of men during the 10 years. Data concerning the other two so-called exercises are of similar nature.

The subtitle of the book, "An Inquiry, Based on Figures Derived From Census Reports, Into The Problems of Ageing Under the Conditions of Modern Industry" is somewhat too presumptuous. The data are distinctly English; not modern industry. The results found by the actuarial manner used by the authors can answer some questions that can be used to advantage for economic planning in England, can help answer such questions as: at what ages are men compelled by reason of their ages to quit the occupations in which they have spent their working lives; what tendency is there in each occupation for earlier or later retirement? Outside of England, the chief contribution of this study is one of method, providing similar and equally reliable data are available and consideration is given to differences in economic and social factors in the industrial scene. Even for England, this study has all the limitations of the actuarial approach unenlightened by situational observations or clinical studies.

Orchard Paper Co. and
Washington University

H. MELTZER

Opinions and Personality. By M. BREWSTER SMITH, JEROME S. BRUNER, and ROBERT W. WHITE. New York, John Wiley & Sons, 1956. Pp. vii, 294.

At the beginning of this book the authors assert: "Opinions . . . are part of man's attempt to meet and master his world. They are an integral part of personality." While the book itself is essentially a monograph which reports case-studies of ten men and their attitudes toward Russia, the case-studies are reported within the framework of a theoretical scheme which stresses the relation between opinions and "general principles of personality, of perceiving, remembering, thinking, learning, motivation." The book reports in detail the case studies of three of the men, while the other seven are described much more briefly. An integrative summary section, however, draws together the material on all ten men and summarizes it in terms of the authors' descriptive scheme.

The description of opinion is divided into three areas: first, a general descrip-

tion of the opinion; second, its expressive nature; and, last, its adjustive functions. The general category refers to an objective analysis of the following aspects: differentiation, salency, time-perspective, informational support, orientation (action-tendencies aroused by the object), object-value (affective tone) and, finally, the policy stand which the individual favors. None of these, the authors emphasize, can be understood apart from the personality of the individual. By the expressive nature of opinion, they mean the dependence of opinion on cognitive functioning, that an individual, for example, who has difficulty with abstractions may have opinions which reflect this capacity. The last part of their scheme, the adjustive functions of an opinion, refers to several motivational functions: that an opinion helps in the evaluation of and response to new information, that the holding of an opinion may aid social adjustment, and that an opinion may represent an externalization of inner needs. The approach to the opinion process does provide a more 'psychological' framework. It is a welcome respite from the usual polling study and should stimulate research.

The methods the authors use deserve comment. Many of them are inspired by H. A. Murray and his co-workers, and have been used previously in the *Clinical Study of Sentiments* or in the *Assessment of Men*. Among the traditional tests employed are the Rorschach, the Wechsler-Bellvue, the TAT, and the Kuder Preference Test. The more Murray-inspired tests include an argument-completion test, a conformity-interview, an information-apperception test, and a stress-interview. In addition, over ten separate interviews were given each S on such topics as childhood memories, family-background, educational history, and personal values. While the theoretical framework helps organize this mass of material and demonstrates the value of considering opinion material in terms of a broader frame of reference, one has the impression that the authors have not gone quite far enough. As it stands the book is an interesting ordering of opinion material in meaningful categories. What one would like, in addition, is a demonstration of the fruitfulness of the scheme. An example might be a study showing that differences in "differentiation" or "time perspective" may lead to different reactions to the same stimulus-situation. Lacking a larger scale investigation, one is given no inkling of which aspects of the descriptive framework are the most productive, which are redundant, and which unnecessary. The report of the ten case studies is not more than a promising start in terms of the theoretical emphasis.

A final comment concerns the writing of the book. In their discussion of the topic the authors have presented a scholarly review of the field, that draws on material from many sources, from personality theory, social anthropology, sociology, social psychology and experimental psychology. This broad perspective is supplemented by clear, interesting exposition that makes the book a pleasure to read. In all, the authors have made a solid contribution to social psychology.

Cornell University

RICHARD D. WALK

Perspectives in Personality Theory. By HENRY P. DAVID and HELMUT VON BRACKEN (Eds.). New York, Basic Books Inc., 1957, Pp. 435.

The origins of this book can be traced to the symposium on European Characterology presented in 1954 at the Fourteenth International Congress of Psychology in Montreal. The volume has the distinction of being the first to be brought out under the sponsorship of the International Union of Scientific Psychology.

The editors have taken the contributions of most of the symposium participants, along with some invited papers, and classified these into three basic sections. The first, *European trends*, describes what is happening in German, Swiss, British, Italian, and French psychological circles. The psychology of personality seems to be flourishing in Britain and Switzerland, growing rapidly in Germany, and developing slowly in France and Italy. In a second section on *Theory*, most space is devoted to elaborating popular European concepts including, Existentialism, Stratification or "levels of mind," character-change, and the "relational" theory of basic needs. There is much here that will be new to the American reader. The third section on *Methodology* is loosely organized. In this set of papers Van Lennep deals with the concept of projection and its particular relationship to the Four-Picture Test; Wellek defends the phenomenological method in psychology (but also asks for a rapprochement with operationism) and Werner Wolff describes an attempt to set depth psychology on experimental ground through his work on expressive movement.

Three very short sections make up most of the remainder of the book. One is called *Commentary* in which Eysenck, with customary vigor and vitriol, criticizes characterology, stratification theory and psychoanalysis on the ground of obscurity, inaccessibility to empirical test, and arbitrariness. Wyatt follows with a plea for a "more systematic methodology of qualitative procedure in psychology" and defends against very narrow definitions of scientific methodology. McClelland, in the final chapter in this section, points out the difficulties in the traditional controversies presented earlier in the book, describes three stages of scientific inquiry (observation, classification, inferences and relationships), bequeaths the practice of making value-judgments about psychology to philosophers, and advocates the building of a science by actively engaging in the three stages of scientific inquiry outlined above.

A final section contains valuable bibliographic material relating to the international picture compiled by H. P. David and a brief historical note on the International Congress by MacLeod.

The book begins with a chapter by Gordon Allport who, after reading the individual chapters, spells out the differences between European and American theories of personality. He sees it essentially as the difference between an emphasis on rigor and an emphasis on ideas. This chapter, delightful reading, serves as an excellent introduction to what follows.

The individual chapters are somewhat uneven. A few are too brief and sketchy to be of any real value in presenting new ideas. Most others are clear and informative. Of those not specifically discussed, the papers of Ellenberger, Franks, Anderson, Frenkel-Brunswik, and Buytendijk are particularly recommended.

Contained within these pages is the flavor of many aged arguments concerning the nature of psychology: *Geisteswissenschaft* vs. *Naturwissenschaft*, Phenomenology vs. Behaviorism, understanding vs. prediction, science vs. metaphysics, and so on. These are arguments which have largely disappeared from the American scene not through any final resolve but rather through lack of communication with our European colleagues during the war years. It is nice to be in touch again. Especially since it seems so obvious that our science is in need of both more rigor and better ideas.

Princeton University

IRVING E. ALEXANDER

Intelligence in the United States: A Survey—with Conclusions for Manpower Utilization in Education and Employment. By JOHN B. MINER. New York, Springer Publishing Co., 1957. Pp. xii, 192. \$4.25.

The data which form the nucleus of this book were collected in connection with another project, but Doctor Miner perceived in them a means of studying the national distribution of intelligence and its relation to several important variables. The principal value that such a survey has for the author is that it provides indications of the waste of native potential with implications for ways of minimizing this waste.

The fact that the sole index of intelligence in the survey was a score on a 20-word multiple-choice vocabulary test is bound to meet with criticism. The author subscribes to the theory of intelligence which holds that there is a single native unit for potential intelligence which is channelized through exposure to various content areas into the development of special abilities. He recognizes that the actual degree of exposure depends upon both the motivation of the individual and the stimulus potential in the environment. It is the wasting of native potential due to limitations in the environmental stimulus with which the author is particularly concerned. Thus, it is a measure of the native potential which he wants, and he devotes the first third of his book to establishing the validity of the vocabulary test for this purpose.

The sample of 1,500 individuals tested was a Public Opinion Surveys sample such as is used in national surveys with stratification in terms of geographical, urban-rural, and economic areas. In addition, a quota system for sex, education, and age (above 10 years) was employed to insure representativeness according to the most recent census data. The adequacy of the sample seems well established.

The second third of the book presents the survey data broken down into the many possible groupings, with tables of means and indices of statistical significance between their differences. The author found that intelligence was most closely related to social stratification (education, race, occupation). Variables of sex, age, area of residence, and religion were less important.

In the last third of this book, the author suggests correctives for inadequate utilization of human potential. He recommends a design of reshuffling of school children and employees so that vocabulary scores within any one grade, regardless of age, and within any one of four "levels" of occupation, will be homogeneous. I am sure that the author cannot be serious about using vocabulary scores as the basis for such an upheaval in the educational and economic systems. But his conviction that the survey indicates inadequate use of manpower and that methods should be explored to correct this is stimulating.

One of the most valuable contributions of the book both to other research workers in the field and to students is the extensive review of the literature at each point of discussion. The list of references numbers 203.

San Francisco

KATHERINE BRADWAY

The Study of Groups. By JOSEPHINE KLEIN. New York, The Humanities Press, 1956. Pp. lx, 200.

This volume possesses unique features which should be brought to the attention of American readers. Its basic objective is to summarize the more pertinent findings

concerning the structure and behavior of groups. It has done this by first examining group behavior in a simplified laboratory situation and bit by bit introducing the reader to increasingly complex situations which approximate more and more the socio-psychological realities outside the laboratory. Each chapter describes a different type of group behavior. The exposition proceeding from a simple situation to the most complex variant through an increasing use of logical analysis and the superimposition of a variety of constraints. This can be illustrated by the manner in which the topic of the "Spread of Information" (Chapter 4) is handled. The 10 assumptions or constraints of this chapter are: let all members be exactly alike; let all members communicate through all the links at their disposal; let the number of members be varied in a group in which each member is directly linked with every other member; let the number of members be held constant and the number of links varied; let the number of links and the number of members be the same in two groups; let the centrality indices of members differ; let there be disturbances in the communication channels; let some members be more closely connected with the network than others; let some members be open to influence and information from the environment; and let the information spread from a single member in a group differentiated into sub-groups.

This volume provides a lucid presentation at the undergraduate level. It does justice to many of the concepts which have become central in the field of group dynamics, concepts such as centrality, communication channel, communication structure, connectivity, hierarchy, leader, transitive and intransitive network patterns, symmetric and asymmetric linkages, formal versus informal communication, etc. These concepts make it possible for the interested undergraduate to push ahead into such more sophisticated areas as those represented by Bavelas' mathematical models for group structure, graph theory and the use of matrix algebra for sociometric analysis. The essentials of the entire subject of organizational structure and its impact upon communication and group behavior are compactly presented.

From this foundation the author makes an easy transition to group behavior closer to the flesh and blood phenomena of everyday life. She provides an ample discussion of the factors responsible for the evolution of intragroup likes and dislikes, group stylization of expressive behavior, communication as adaptive behavior, group factors which mediate decision making, the relationship of the group processes she has described to society taken as a social system, and finally the group phenomena which inaugurate and maintain social change.

The book appears to have been influenced chiefly by the ideas of Bales and Homans. It should be of genuine interest to psychologists, sociologists and methodologists.

Hollins College

HENRY WINTHROP

Freudiana: Including Unpublished Letters from Freud, Havelock Ellis, Pavlov, Bernard Shaw, Romain Rolland, et Alii. By A. A. ROBACK. Cambridge, Mass., Sci-Art Publishers, 1957.

The title of this little book has a certain appropriateness but this brief notice will deal only with the stricter contribution made by these pages to the biography of Sigmund Freud. Between the years 1930 and 1939 (the year of Freud's death), Roback received 19 letters from Freud. In addition to a one-sentence introductory

letter (that apparently is not supposed to count) six longer ones are translated in this book, the original German of them being also given in an Appendix. (The author informs us that he can reproduce six only because it is the policy of the Sigmund Freud Copyrights Ltd. "not to countenance the publication of a complete correspondence with any individual, except in the case of the definitive biography.") In these letters Freud comments on some of Roback's writings and responds negatively to an abortive plan of the latter to publish a Freud "jubilee book." In addition, there is reproduced a memorial leaf from the Freud family Bible that was presented to Sigmund by his father with an added inscription on his son's thirty-fifth birthday in 1891. A translation of the inscription is given. There are 13 illustrations in the volume, three being of letters in Freud's autograph. Since it is not intended to offer a conventional review, let it suffice merely to mention a minor discrepancy: the photograph on page 206 is described in the List of Illustrations as being of the "Clark University Centennial Group with Freud and Associates"; on page 206 itself, under the picture, we read "Galaxy Attending Decennial of Clark University, September 1909." The celebration in question was to commemorate a *twentieth* anniversary.

The letters, constituting the durable contribution of the volume, have the recognizable flavor of Freud's personality. The tone becomes more friendly as the correspondence progresses, but perhaps the most significant (and trenchant) of the letters is the first of the longer ones. It reads, in part, as follows:

"Today I received your letter as well as a small library of your writings and don't want to delay thanking you for them until I have read them. Besides, I feel impelled to put a stop to a situation that is very unusual for me. For over 30 years I have let people say and write whatever they wish about myself and my teachings. Only in very exceptional cases do I make a rebuttal. (e.g. in *History of the Psychoanalytic Movement*). I know very well that my protests would not do any good, since people feel a need for expressing themselves in just the way they do. And now I find myself involved in an exchange of criticism and counter-criticism with you, although I cannot dispute *your* right to express yourself about me as freely and incorrectly as you please. I know this exception is due to the fact that you strike a chord of Jewishness which reverberates so sensitively in me. My sympathies were aroused, and then I was sorry that I seemed to find a discrepancy between the high position you wish to accord me and your knowledge about me and your understanding of my work. No harm meant; I won't do this sort of thing again.

"I am glad to hear that you yourself take criticism well. I shall give you an opportunity for that by making a few remarks regarding myself and psychoanalysts."

Washington University

SAUL ROSENZWEIG

Epilepsy: Grand Mal, Petit Mal Convulsions. By LETITIA FAIRFIELD. New York, Philosophical Library, 1957, Pp. 159. \$4.75.

This excellent little volume is deceptively unpretentious. It will help to dispel many misconceptions about epilepsy that are still distressingly widespread and deep rooted for it is lucidly written, accurate, astonishingly complete and up to date.

The basic approach is common-sense and the language nontechnical. Those things are said about epilepsy that are agreed upon by the experts, but this is not a

purely derivative work; much of it is based on first-hand experience. Even the expert will find the manner and the matter interesting. The greatest value of this book will be as an introduction to epilepsy for psychologists, physicians, teachers and others who have a special concern with this disorder. Intelligent parents of epileptics and intelligent adult patients will also find it useful for they will discover here relief from many perplexities. Controversial material is omitted, and this is easily justified; new data developed from recent research must be validated by competent authorities before being presented to practicing physicians or the interested public.

Some American experts will regret that the relationship between psychomotor epilepsy and personality defects has not been stressed and the frequency with which certain types of behavior disturbances occur in association with 14 and 6 per second positive spikes is not mentioned. Until these facts are recognized, the generality of epileptics will have to bear the bad name which properly belongs to certain rather specific types of epilepsy. In this book infantile spasms are not identified as a form of epilepsy characterized by mild clinical seizures, maximal electroencephalographic abnormality and maximal risk of mental retardation. Nor are repeated febrile convulsions recognized as a special entity with a high hereditary factor and an excellent prognosis. The darkest and lightest spots of epilepsy are not shown in their true values and some important details are lacking, but this book gives a clear and enlightening picture of a common and much misunderstood disorder.

The sections that deal with clinical care and custodial facilities for epileptics in England are not entirely pertinent to America and one might wish that the American edition had been modified to give the equivalent information for the United States and Canada. However, the 48 states have no uniform provisions for epileptics so this might be impractical.

Every chapter contains useful information. Even the parts that are essentially British (the chapters on *Epileptics in the Community*, *Colonies for Epileptics*, *Epileptics and the Law*) are interesting. Appendix 1 tells the story of the British Epilepsy Association. Appendix 2 is a statistical table indicating the size and character of the employment problem created by epilepsy; these are statistics that are hard to come by and are most welcome in this handy form. Appendix 3 contains the 40 most frequently asked questions about epilepsy and their answers. There is an excellent brief, highly selective bibliography. The lack of an index is no serious disadvantage in a short book like this; chapters have simple descriptive titles and are separated into sections and paragraphs with italicized headings.

University of Illinois College of Medicine

FREDERICK A. GIBBS

When Prophecy Fails. By LEON FESTINGER, HENRY W. RIECKEN, and STANLEY SCHACHTER. Minneapolis, Minn. University of Minnesota Press, 1956, pp. v, 256.

This is more than a fascinating account of an end of the world movement that briefly flourished in our midwest a few years ago. For, in addition to a closeup view of the fantastic persons and events that comprise the life and death of the movement, the authors—who infiltrated to the inner circle pretending to be loyal believers—report a field test of a single hypothesis derived from Festinger's theory of "dissonance". The opportunity for such a test arose because, unlike most millen-

nial or messianic movements, this one specified the time, place and mode of salvation. (A space ship would land in the backyard of the leader and carry the chosen group off to a safe planet.) Thus, when neither the deluge nor the space ship appeared, it was possible to observe the reactions of the followers to a disconfirmation of at least a great portion of their belief system.

Briefly, the hypothesis put to test states that, given a deep commitment to a belief that is shared with others, a clearly perceived disconfirmation results in increased proselyting in behalf of the belief. Theoretically, proselyting is reviewed as a means of strengthening the belief through the addition of "social reality". By developing a stronger belief than before the disconfirming event, the follower is able to bring about a net reduction in the dissonance created by the disconfirmation. Detailed anecdotal observations show in fact that proselyting does increase, at least on the part of the central members of the movement. Thus, interviews were freely given to the press, broadcasts were made, etc.

As is clearly recognized by the authors, the hypothesis is difficult to affirm unequivocally by the study of a single case. There would seem to be alternative explanations. For instance, assuming that disconfirmation was privately acknowledged, increased proselyting could have been due to a desire of the members to give each other comfort and support in the wake of an ideological disaster. That is, while privately discarding their invalidated beliefs, everyone put on a show of unshaken faith in the cause to bolster each other's fast-sinking morale. If the reviewer reads the theory correctly, the increased proselyting should, strictly speaking, be due wholly to the attempt to strengthen a potentially weakened belief and not due to an attempt to make the painful disappointment more tolerable for one's comrades. But these are hairs whose precise splitting will have to be done by laboratory experiments.

This is clearly a bold and imaginative piece of research that will interest and instruct psychologists regardless of their theoretical predilections. And, for those who go in for real (sic!) science-fiction, the book is a must.

University of Pennsylvania

ALBERT PEPITONE

Biochemical Individuality, The Basis for the Genetotropic Concept. By ROGER J. WILLIAMS. New York, John Wiley and Sons, Inc., 1956, Pp. xi, 214.

In the first eleven chapters, the author has assembled data indicating the wide variations among individuals of different species with most attention given to man. In the section devoted to anatomy, illustrations are reproduced from Anson's *Atlas of Human Anatomy*. Other chapters have assembled data from the realm of biochemistry and treat of enzymes, tissue composition and metabolism. A whole chapter is devoted to nutrition and variations in nutritional requirements. The need for the consideration of genetics in all of these relationships is stressed in a manner that would hardly be worth stating to a geneticist but is probably novel to many biochemists. No attention is given to the ever-increasing individuality of man as he grows older and pathological changes accentuate earlier differences.

Each chapter has bibliographies that should be very helpful to those wishing to read more about individual variations.

The last three chapters are entitled "implications" for biology, medicine, and psychiatry. The last chapter may prove disappointing to those trained in psychology.

It repeats the ideas of the last century that researchers should be able to detect variations in the structure of the brain that would account for individual differences in function. It fails to mention the life-time pursuit and frustrations of some of the specialists who have labored with collections of human brains such as that assembled at Cornell. It states that there must be a physical-chemical basis for mental diseases but does not stress that chemical methods of measurement have proved inadequate. Some discussion of the author's well known studies of alcoholics is included.

This interesting book should prove very worthwhile reading for biochemists, trained primarily in the physical sciences. To biologists the ideas will be more familiar. To physicians, used to dealing with problems resulting from the individuality of patients, there will be little new in generalizations indicating that patients vary in reactions to drugs, disease, and environment. The data included in the various chapters are not, however, extensive enough to provide a handbook of variability. Since the days of Noah, animal industries have based their operations upon the individual performance of animals. The individual cow is culled from the flock because she does not produce adequate milk. Therefore the philosophy of the book will be well understood by animal husbandrymen but the work gives almost no attention to the individuality of farm animals.

There is a brief subject index.

Cornell University

C. M. McCAY

The Organization of the Cerebral Cortex. By D. A. SHOLL, London, Methuen & Co. New York, J. Wiley & Sons, 1956. Pp. 125.

The first part of the book gives a brief review of the older facts and concepts of cortical anatomy and physiology. Some of these the author considers as inadequate. He has made his own contribution in presenting a study of the cellular patterns of the cerebral cortex with special reference to the total number of neurons, their packing density, and the functional connections of large neuronal assemblies as functional units. His main concern is with neurons in large aggregates whose axons presumably have a definite destination to other areas of the cortex nearby, or to subcortical dependencies. The author holds that histological studies in the past show that no theory, mathematical or other, relying on specific neuronal circuits can be maintained. He regards the multiplicity of radial fibers, nerve cells, their axons and collaterals as the elements which in the aggregate conduct various impulse complexes through the cortex. He was able to validate this by quantitative histological studies of these aggregates of cells and their axonal collaterals. However, the destination of their axons in the aggregate to other areas or to subcortical structures is not shown by this method alone. Horizontal contacts across cells are indicated. The analysis of the horizontal bands or stripes has not progressed far; nor is it possible by quantitative histological methods to do this effectively for a large number of axons. One can, however, accept this as a probable outcome.

The author supports the idea, as does the reviewer, that large assemblies of neurons are in action in the cortex and create dynamic domains. Presumably these domains have also subcortical repercussions. For "if one domain is brought under the presence of some external charge, then the pattern of orientation of that domain

may be changed, and a wave of activity may pass through the whole assembly giving rise to a new pattern which may be very complex. Such a theory (of Cragg and Temperly, 1954) is far in advance of its predecessors in its agreement with known anatomical and physiological features of cortical organization." This is another version of mass action which admits of a large number of units acting in concert.

The usual cortical areas, their transcortical and subcortical connections receive less consideration in this discourse. The stellate cells and granule cells are not assigned a special part. Cholinesterase as a fluid which conducts nerve impulses, and the modalities on such multiple networks are not introduced. Chemical resonance as a means of recording is implied by allusion to hysteresis. Many contributive works are discussed, and most of the arguments are well presented. The reviewer believes that this approach, with more physical and chemical structures added to the solution of the problem of field activities of the cerebral cortex, should create an interest for itself in the future.

Columbus State Hospital

JAMES W. PAPEZ

Fads and Foibles in Modern Sociology and Related Sciences. By PITIRIM A. SOROKIN. Chicago, Henry Regnery Company, 1956. Pp. viii, 357.

There seems never to be an end to the literary output of this testy Russian, and (let's face it) he has plenty to say. Never a pussyfooter, he lays about him with mighty strokes, naming names as he lines up his opponents, the better to sever many heads with each blow.

In this volume he takes contemporary sociologists and psychologists to task, often deservedly, for their many and varied sins. First to draw his fire are claims of new discoveries for ideas and insights long in the literature, and failure to give credit for concepts obviously borrowed from others. He doubts that sloppy scholarship and unconscious forgetting can explain more than a small fraction of such infringements. At the same time he takes a deserved slap at the cliques that develop in particular areas of research whose members give generous recognition to their fellows, ignoring completely similar, earlier, or even better work by others. I applaud with enthusiasm when he pillories the writers and scientists with nothing new to contribute but neologisms and an esoteric terminology, and when he attacks the numerous attempts to develop elaborate "operation" formulations of sociological and psychological problems and to quantify where there are no units of measure.

Then the testers of all sorts come in for a drubbing, but here our sociologists has missed the boat, for he damns, as far as I can see, everything ever done in the entire field. One can only conclude from his furious assault that he assumes psychological tests to be impossible in principle, and that is laying it on a bit thick. Of course testing has its troubles, limits, and pitfalls. It has its naïve enthusiasts, and, by reaction, its equally naïve denouncers, but even if we should grant that all tests today in existence are worthless, which this reviewer certainly does not, that would then be the very best reason in the world for developing some good ones. The need and the demand for the kind of data that the tests seek to supply long ago produced such aberrant manifestations as astrology, palmistry, and physiognomy. Should we have let things drop there? If the problem is not inherently

insoluble, where are we to look for a method that might bear fruit? Not, according to Sorokin, in the entire area of "physiocalistic" psychology, under which rubric he includes everything from the conditioned response techniques of behavior study to somatic types and dianetics. For, according to Sorokin, causal determinism just does not operate in the area of the psycho-social; volition is not a deterministic process. It is no surprise then to find him insisting upon "supersensory reality," and non-sensory or supersensory modes of cognition. In short, we have here intuitionism and mysticism; a hardy perennial. If enough of these mystics turn up in any of these perennially touted "swings to mysticism" they should be carefully studied with the aid of some of the products of the testing laboratories for which Sorokin has such Olympian contempt. There'll be gold in them thar personality profiles!

New York University

GEORGE B. VETTER

Backwardness in Reading: A Study of Its Nature and Origin. By M. D. VERNON. New York (also London), Cambridge University Press, 1957. Pp. viii, 228. \$4.75.

In the past, systematic treatises in the field of reading disability have been few in Great Britain as compared with America. Consequently the present work fills an important need for British psychologists and educationalists. Also it contains much important information for interested Americans.

The main sections of this book are concerned with visual perception in reading, auditory perception in reading, nature of reading disability, innate factors as causes of disability in reading, relation of various acquired defects to reading disability, effect of environmental factors on reading ability and disability, and cure of reading disability. The appendixes contain case studies and citation of some supplementary data. There is a list of references and a good index. In general the book is concerned with the more severe types of reading disability. Heavy emphasis is placed upon investigation of the individual case.

Several desirable aspects of the treatise deserve special mention: (1) With few exceptions the experimental literature is covered adequately. (2) The discussions are well organized and well written. (3) Most of the views on the nature of the reading process, on reading disability and on visual perception of printed words are sound. (4) The author is justifiably critical of data derived from the use of unsatisfactory experimental designs. (5) Evaluation of the use of phonetic analysis is very well done. (6) Proper emphasis is placed upon verbal facility as a factor in learning to read. (7) The relation of visual and auditory defects to reading disability are handled satisfactorily. This is also true for such environmental factors as home environment and teaching efficiency.

A number of criticisms of this treatise may be cited: (1) Although multiple causation of reading disability is mentioned, unfortunately it receives too little emphasis. (2) It seems to the reviewer that the author is overcautious in her evaluations. Repeatedly she cautions the reader to await further experimental evidence, or that a certain effect *may* have occurred, or to be aware of contradictory findings. A more positive emphasis upon trends derived from the more adequate experiments would have been a good supplement. Too frequently the reader gets the impression that contradictory findings mean that nothing of value has been achieved. (3) Apparently the author attempted to cover any and all literature that *might be* relevant. At times (as with cerebral dominance, for instance) this has

led to a bewildering citation of data, some of which are of doubtful significance. (4) The role of intelligence in reading disability is inadequately treated. Particularly, no emphasis is placed upon the fact that children with low intelligence drop behind the rest of the class and this eventually results in reading disability. (5) Too little emphasis is placed upon maladjustment due to reading disability. This has resulted in an overemphasis on emotional disorders as causes. As a matter of fact, throughout the book there is too great a stress on innate factors as causes. Actually most reading cases are made. They are not born that way. (6) The chapter on the cure of reading disability is especially inadequate. Experimental results are cited to show that improvement in reading occurs with remedial teaching. However, no details for diagnosis and remedial instructions are given. (7) The author considers that reading disability may be due to some fundamental cognitive incapacity, as yet unidentified. She has little sympathy with American methods of diagnosis and remediation even though they are successful. (8) The case studies are inadequate as to details of diagnosis and remedial treatment. (9) Americans will be annoyed by the incompleteness in citation of references. Titles of articles are omitted and page numbers are incomplete.

In general, this book should be favorably received by those dealing with reading disability. It is the best British publication on the subject. The good points outweigh the shortcomings by far. And the reader should not forget that the author is concerned mainly with severe disability.

University of Minnesota

MILES A. TINKER

American Families: A Demographic Analysis of Census Data on American Families at Mid-Century, By PAUL C. GLICK. New York, John Wiley and Sons, Inc., 1957. Pp. xiv, 240.

One of a series of Census Monographs sponsored by the Bureau of the Census and the Social Science Research Council, this study presents an analysis of demographic data on family life in the United States during the 1940's and early 1950's. These monographs aim to make information secured by the censuses fully available to satisfy the needs and interests of the community and to constitute a broad base for further studies in the social sciences. Dr. Glick, who is chief of the Social Statistics Branch of the Bureau of the Census, treats his subject-matter with a restrained enthusiasm and writes in a clear and lively style. He states that his hope is that the descriptive material presented may serve as a basis for the testing of theoretical hypotheses, and that it will be useful to those persons who depend upon such reliable, nationwide, family-statistics to answer their questions.

Adults in the age range between but not including adolescence and old age are the center of much of the discussion. It concerns itself with the fields of family-composition, social and economic characteristics of the family, family formation and dissolution, and future household formation. In a description of the life cycle of families, for example, there is a demographic analysis of characteristics and changes which families go through between formation and dissolution. Here, new basic data relate to the advancing age of the family head. Covered also are such subjects as the employment cycle for the man, economic status, cycle of labor force participation of the wife, and residential mobility. Differential patterns in family behavior among the several social groups are delineated. The author, too, has traced historical

trends and has commented upon the implications of the findings for future population-changes. Numerous tables and graphs form a part of the text materials. Completing the volume are helpful appendices.

Child Welfare Services
Los Angeles, California

MARJORIE B. DRURY

College Teaching: A Psychologist's View. By CLAUDE E. BUXTON. New York, Harcourt, Brace and Company, 1956. Pp. ix, 404.

The title of this new book succinctly describes its contents. The fourteen chapters fall roughly into four areas. Introductory chapters deal with higher education, its history, basic issues, and the place of psychology in the curriculum. Roles, responsibilities and decisions in the career of the college teacher are considered next. Third, making up the core of the book, is a series of chapters on the instructional process: determining course objectives, lecturing, leading discussions, trying newer methods, and constructing and scoring examinations. Finally, there is a series of discussions of knotty student-faculty problems such as class morale and advising students.

For the reviewer, there were several important highlights. Throughout, the book is well organized and carefully written. The author proceeds logically and systematically with deceptive clarity. Second, in his chapter dealing with the typical undergraduate training program in psychology, the author criticizes, with considerable penetration, features such as unstated objectives, overlapping of courses and lack of sequence in subject matter. This is not psychology's problem alone, as is indicated. Best of all, particularly in the teaching-learning discussions, is the impression one receives of a first rate applied psychologist in education. Here is a person versed in learning-psychology, possessed of many years of teaching experience, and schooled in the available literature. His comments raise many problems, solve a few, and should provoke continued research and thought about others.

This reviewer noted two features that might be considered omissions. There was little extrapolation from the research that has been done at the secondary level. Some of the better material in the sciences and social studies would have contributed to the author's analysis of college teaching. Secondly, it would have been of interest to have had a summary of learning theory which would integrate the scattered findings in the area of how and why college students learn. Does not psychological theory have a contribution to make to instructional theory?

The book should prove of value to many: to the new instructor, to the teacher in seminars on teaching at the college level, as a fine text, and even to teachers of many years' experience, who may achieve additional insights. This is an important view of college teaching by a thoughtful psychologist.

Washington University

LOUIS M. SMITH

INDEX

CHARLES VALLEY BROOK, University of Texas

AUTHORS

(The names of authors of original articles are printed in CAPITALS; of authors of books reviewed, in roman; and of reviewers, in *italics*.)

ADAMS, J. K.	432	CLARK, R. E.	97	GLEITMAN, H.	117
ADAMS, P. A.	642	CLAUSEN, J.	577	Glick, P. C.	676
<i>Alexander, I. E.</i> ...	151, 664	COFER, C. N.	606	GOTTSANKER, R. M.	110
Ambrose, G.	661	COHEN, W.	403	GRANGER, G. W. ...	462
ANDERSON, L.	411	COLE, D. L.	654	<i>Gretther, W. F.</i> ...	658
Ansbacher, H. L. ...	151	CONKLIN, J. E. 132,	395	GRUBER, H. E.	133
ATTNEAVE, F.	318	CROCKETT, W. H. ...	283	GUILFORD, J. P. ...	427
		Cruickshank, W.	153	GULICK, W. L.	276
BAKER, K. E. ... 81,	268			GUTTMANN, H. E.	453, 456
<i>Baker, K. E.</i>	329	DALLENBACH, K. M.			
BATTIG, W. F.	75	 135, 471,	655	Hagenah, T.	325
BENDIG, A. W.	286	Darley, J. G.	325	<i>Harper, R. S.</i> ... 159,	663
BERNARDIN, A. C. ...	133	David, H. P.	664	Harris, R. E.	145
BILODEAU, E. A. 49,	550	DAVIS, R. T.	87	HASLERUD, G. M. ...	97
BITTERMAN, M. E.		DAVITZ, J. R.	69	HASTORF, A. H. ...	114
..... 135,	569	DAY, W. F.	238	HEIMER, W.	493
<i>Bitterman, M. E.</i> ...	141	DEESE, J.	594	<i>Held, R.</i>	331
Blair, G. W. S.	329	<i>Di Michael, S. G.</i> ...	659	Henry, W. E.	158
Blum, M. L.	159	DINNERSTEIN, D. ...	21	HERON, W.	38
Bond, G. L.	332	<i>Doll, E. A.</i>	153	HESSERT, R. B.	628
<i>Bradway, K.</i>	668	Dorcus, R. M.	661	HILGARD, E. R. ...	472
BRANCA, A. A.	541	<i>Dreher, R. E.</i>	155	<i>Hochberg, J. E.</i> ...	480
BRAUN, H. W.	628	<i>Drury, M. B.</i>	676	HOHLE, R.	272
BRENNAN, G.	248	DUDEK, F. J. 81,	268	HUNT, W. A.	347
BROGDEN, W. J. ...	75	DUKES, W. F.	306	HURVICH, M. S. ...	258
Brown, C. W.	155	DUNCAN, C. P. ...	616		
BROWN, R. H.	376	Dunne, A. C.	667	Ikin, A. G.	487
BROWN, W. L.	126				
BROWNSTEIN, A. J. ...	123	EDWARDS, R. V. ...	110	JAHNKE, J. C.	319
<i>Brózek, J.</i>	149			JENKINS, W. L. ...	640
Bruner, J. S.	665	Fairfield, L.	670	JOHNSON, C. W. ...	64
Brunswick, E.	480	FALK, J. L.	303	JONES, F. N.	227
BUCHWALD, A. M. ...	233	Farnsworth, P. R. ...	323	<i>Jones, F. N.</i>	323
BURNHAM, R. W. ...	386	<i>Feshbach, S.</i>	147	<i>Jones, M. H.</i>	323
Burton, A.	145	Festinger, L.	671	Jones, M. R.	147
BURT, H. E.	656	FISKE, D. W.	469		
Buxton, C. E.	677	FORD, T. J.	606	Kaplan, O. J.	156
		FORREST, D. W. ...	634	KARLIN, L.	248
CAMPBELL, D. T. ...	347	FRANKS, C. M.	462	KARR, A. C.	640
CAPALDI, E. J.	451	FRICK, J. W.	427	KASWAN, J.	447
<i>Carpenter, C. R.</i> ...	156			Katz, E.	157
Carroll, J. B.	152	GARDNER, M. L. ...	299	KENNEDY, J. L. ...	114
CHAPANIS, A.	650	GARDNER, R. A. ...	174	Klein, J.	668
Cherry, C.	322	Ghiselli, E. E.	155	KLEMMER, E. T. ...	56
CHO, J. B.	87	<i>Gibbs, F. A.</i>	670	KOLERS, P. A.	437
Clark, F. L.	667	GIBSON, J. J. 64, 129,	291		

<i>Landis, C.</i>	156	OVERTON, R. K.	126	STAATS, A. W.	442
<i>LANE, G.</i>	468			STAGNER, R.	299
<i>Lane, H. S.</i>	663	<i>PAPEZ, J. W.</i>	333	STAKE, R. E.	276
<i>LANGHORNE, M. C.</i>	321	<i>Papez, J. W.</i>	673	STANLEY, J. C.	467
<i>LAWSON, R.</i>	123	<i>PAPEZ, P.</i>	333	STEINSCHNEIDER, A.	64
<i>Lazarsfeld, P. F.</i>	157	<i>PARDUCCI, A.</i>	194, 272	STEVENS, J. C.	600
<i>Lebo, D.</i>	487	<i>PASTORE, N.</i>	131	STEWART, P. A.	528
<i>LEIBOWITZ, H.</i>	106	<i>Pepitone, A.</i>	158, 671		
<i>Levine, E. S.</i>	663	<i>POFFENBERGER, A. T.</i>		TALLAND, G. A.	263
<i>LEWIS, N. A.</i>	347		136	THOMAS, G. J.	528
<i>Lofquist, L. H.</i>	659	<i>PORTER, L. W.</i>	219, 421	Tinker, M. A.	332
<i>LOUTTIT, C. M.</i>	314	<i>POSTMAN, L.</i>	465, 642	<i>Tinker, M. A.</i>	675
		<i>POULTON, E. C.</i>	631	THURLOW, W.	653
<i>MAAG, C. H.</i>	243	<i>PRESSEY, S. L.</i>	656	TULVING, E.	600
<i>Marcuse, F. L.</i>	661	<i>PRONKO, N. H.</i>	128	Tyler, L. E.	149
<i>MARDER, V. L.</i>	594				
<i>MASON, D. J.</i>	69	<i>RAY, W. S.</i>	289	Underwood, B. J.	658
<i>McCay, C. M.</i>	672	<i>RIDGLEY, M.</i>	303		
<i>McGILL, T. E.</i>	637	<i>Riecken, H. W.</i>	671	<i>VANDERBILT, C.</i>	577
<i>McNemar, Q.</i>	323	<i>Roback, A. A.</i>	669	Van Pelt, S. J.	661
<i>Meltzer, H.</i>	667	<i>ROBERTS, H. C.</i>	299	VAUGHAN, C. J.	286
<i>MELVILLE, J. R.</i>	316	<i>ROCK, I.</i>	186, 493	VERNON, J. A.	637
<i>MESSICK, S. J.</i>	161, 586	<i>ROCKLYN, E. H.</i>	628	Vernon, M. D.	675
<i>MEYER, M. F.</i>	203, 646	<i>Roe, A.</i>	663	<i>Vetter, G. B.</i>	674
<i>MICHAEL, A. L.</i>	299	<i>ROSEN, E.</i>	120	VIEK, P.	69
<i>Michael, W. B.</i>	330	<i>Rosenzweig, S.</i>	669	Von Bracken, H.	664
<i>Michaelis, A. R.</i>	156	<i>RUSSELL, W. A.</i>	120	Voss, J. F.	75
<i>MILLER, G. A.</i>	311	<i>Ryan, T. A.</i>	325		
<i>Miller, G. A.</i>	152, 322	<i>SALTZMAN, I. J.</i>	253	<i>Walk, R. D.</i>	665
<i>Miner, J. B.</i>	668	<i>SARASON, B. R.</i>	92	<i>Ware, K.</i>	332
<i>MITCHELL, J. M.</i>	459	<i>SARASON, I. G.</i>	92, 620	WATSON, W. A.	560
<i>MONCRIEFF, R. W.</i>	1	<i>Sarbin, T. E.</i>	158	WAYNER, M.	457
<i>MORSE, W. H.</i>	308	<i>Schachter, S.</i>	671	WEHRENBURG, L. JR.	128
<i>MOWRER, O. H.</i>	69	<i>Schilder, P.</i>	661	WEINSTEIN, S.	611
<i>Muenzinger, K. F.</i>	154	<i>SHAKLEE, A. B.</i>	624	WEISSMAN, S. L.	283
<i>MULHOLLAND, T.</i>	280	<i>SHERLOCK, L.</i>	102	WERTHEIMER, M.	21
<i>Murphy, L. B.</i>	327	<i>SHIPLEY, T. E. JR.</i>	258	White, R. W.	665
<i>Mussen, P.</i>	327	<i>Sholl, D. A.</i>	673	Williams, R. J.	672
		<i>SKINNER, B. F.</i>	308	<i>Winthrop, H.</i>	668
<i>NACHMAN, M.</i>	211	<i>Smith, L. M.</i>	677	WISHNER, J.	258
<i>NAGEL, E. H.</i>	75	<i>Smith, M. B.</i>	665	<i>Wishner, J.</i>	145
<i>NEFF, W. D.</i>	320	<i>SMITH, O. W.</i>		WODINSKY, J.	569
<i>NEISSER, U.</i>	512		64, 102, 291, 361	WOLFF, W.	470
<i>Newbold, G.</i>	661	<i>Snyder, L. L.</i>	154		
<i>NEWMAN, E. B.</i> 135, 652		<i>SOLLEY, C. M.</i>	161, 586	<i>Yagi, B.</i>	485
		<i>Sorokin, P. A.</i>	674	<i>Yatabe, T.</i>	485
<i>O'KELLY, L. I.</i>	303	<i>Spence, K. W.</i>	141	YOUNG, P. T.	295
<i>OSGOOD, C. E.</i>	411			ZEIGLER, H. P.	106

SUBJECTS

(References in *italic* figures are to reviews.)

- Adaptation, dark, eyelid reflex during, 462 ff.; GSR, 462 ff.; smell, 1-20.
- After-effect, associative, perceptual organization, 490; color, interocular, 318 f.; figural, group-method for study of, 319 f.
- After-images, apparent size, 114-116.
- Aging, chronological vs. psychological, 459 ff.; industry, in, 664 f.; psychological, an hypothesis of, 459 ff.
- Animal: Bestiary, 20th century, 311 f.
- Chicks, responses to cues of depth, 132.
- Fish, discrimination-reversal in, 569-576.
- Monkey, imaginary, random movements at typewriter, 311-314; preferences, 87-91.
- Pigeon, superstition in, 308-311.
- Planaria, maze for, 128.
- Rats, conditioning of fear, 69-74; cutaneous electrical thresholds, 624-627; resistance to extinction, 123-125.
- Research, cutaneous electrical thresholds, 624-627.
- Small, apparatus for record of fluid intake of, 295-298; respirator for, 457 f.
- Thinking in, 485 f.
- Announcements, editorial, addition to editorial board, 135; book-review editor, 135.
- Anxiety, manifest, 286-288.
- Apparatus: Curvature judgments, for, 361 f.
- Device for production uniform illumination of entire visual field, 404 f.; lever-pressing, for fluid rewards, 303-305.
- Discriminative learning, for fish, 570 f.; reaction-time, 57.
- Drinkometer, for small animals, 295-298.
- Maze, for Planaria, 128; variable pattern, inexpensive, 306 f.
- Motion, translatory, 103, 291-294; visual translatory, 291-294.
- Movement, perception of, 65 f.
- Odorometer, 1.
- Olfactometer, 1.
- Problems, verbal vs. manipulatory solutions, 289 f.
- Pursuit-test, multidimensional, 551 f.
- Pursuit-tracking, 49 f., 75 f.
- Respirator, inexpensive, for small animals, 457 f.
- Sound-key, circuit for, 456 f.
- Stimulator, constant current sine wave circuit for electrical, 453 ff.
- Tachistoscope, 38-48, 97-101, 299-302; three-field Dodge, 438; use of, in differential recognition of words, 316-318.
- Tracheal cannula, 457 f.
- Tracking, 49 f.
- Tracking-box, 110.
- Visual, for 'empty-field' myopia, 378 f.
- Visual beats, 578.
- Voice-key, circuit for, 456 f.
- Assimilation, contrast and central tendency, 560-568; judgments of clinical materials and, 347-360; proactive and retroactive, 117-119.
- Association:
- Contingency content, analysis of, 411-420.
- Fittingness and intent to learn, 447-450.
- Free, communality of, 120-122.
- Hypotheses, six theoretical, 411-420.
- Law of effect, 219-226; recent studies of, critically reviewed, 642-648.
- Nonsense figures, of, 447-450.
- Reaction-time, free, verbal context and, 606-610.
- Role of repetition and insight in, 186-193.
- Theory, assumptions and implications tested, 411-420.
- Word, 597 ff.; auditory vs. visual presentation, 233-237; free, frequency characteristics of, 120-122; norms of, 233-237.
- Attention, perception, a function of, 38-48; recognition and, 38-48; retinal fixation and, 38-48.
- Aubert-Fleischel paradox, 64 f.
- Audition: Autokinetic effect in, 133 f.
- Deafness, congenital, human development and, 663 f.
- Frequency-span of human voices, 647.

- Hydraulic theory, 205.
- Loudness, estimations of, 600-605; successive comparison of, 117-119; Tartini and beat-tones pitches, 646-649; Tartini pitches, created by prime number seven, 203-210.
- Auditory, figure-ground effect, 653 f.
- Autokinesis, auditory, 133 f.
- Backwardness, reading, in, 675 f.
- Beats, Tartini, 203-210, 646-649; tones, 648-649; visual, 577-585.
- Behavior, theory and conditioning, 141 ff.
- Belief, prophecy and, 621 f.
- Bestiary, 20th century, 331 f.
- Betting, two-choice discrimination in, 195-202.
- Biography: Goddard, H. H., 656 f.
Hollingworth, H. L., 136-140.
Terman, L. M., 472-479; portrait, facing 233, acknowledgment of, 471.
- Brain, cortical integration, field-theory of, 528-540; stimulation of, effect on visual perception, 528-540.
- Chicks, *see* Animal.
- Children, exceptional, symposium on, 153 f.; perception, size, 106-109; reading difficulties of, 332; young, personality of, 327 ff.
- Choice-responses, objects for (monkey), 87-91; rate of making, factors affecting, 650 ff.; comment on, 652 f.
- Clinical psychology, mental disorders in later life, 156; personality, 145 ff.
- Cognition, recognition and, mathematical theory of, 322 f.
- College, teaching, 677.
- Color, after-effects, interocular, 318 f.; differences, inter-eye, 386-394, 655; weakness, compensatory learning and, 126-127.
- Communication, human, mathematical theory of, 322 f.
- Concept, formation, 586-593; transposition, 263-267.
- Concepts, meaning and statistical structure of, 161-173.
- Conditioning, fear, 69-74; review of, 141 ff.; semantic generalization and, 541-549.
- Confidence, scale of, 432-436.
- Constant-sum method, modification of the, 81-86, 268-271.
- Context, verbal, free association-time and, 606-610.
- Contrast, assimilation and central tendency, 560-568; judgments of clinical materials and, 347-360.
- Convulsions, epileptic, 670 f.
- Cortex, organization of, 673 f.
- Cortical integration, field-theory of, 528-540.
- Counseling, vocational, physically handicapped, 659 f.
- Critical flicker frequency, brain stimulation and, 531 ff.
- Cues, air-pilots, for landing, 130.
- Cutaneous, sensation, paradoxical warmth, 640 f.; thresholds, electrical (rat), 624-627.
- Deafness, congenital, human development and, 663 f.
- Decisions, complex, rate of making, factors affecting, 650 ff.; comment on, 652 f.
- Deficiency: Mental, brain damage in, 333-346.
Cerebral changes in, 340-343.
Cytopathological changes in, 334 f.
Dementia praecox and, 345 f.
Hypothalamic changes in, 343-345.
Mycotic changes in, 333-340.
Schizophrenic responses, 348 ff.
- Depth, perception of, without textural gradients, 611-615.
- Differences, color, inter-eye, 386-394, 655; sex, abstract reasoning, 446.
- Differential, thresholds, visual intensity, 211-218.
- Differential psychology, 149 f.; group differences, 150; individual differences, 149, 151.
- Discrimination: Categories, absolute and relational, 263-267.
Criteria of, 263-267.
Fish, in, 569-576.
Intensity, visual, 211-218.
Reinforcement ratios, of, in betting, 194-202.
Size, anxiety and, 286-288.
- Distance, apparent visual size and, 106-109.
- Disturbances, rhythmic, effect of, 56-63.
- Drive, manipulatory (monkey), 87.
- Eastern Psychological Association, *see* Meetings.
- Editorial announcements, 135.
- Effort, effect of, on extinction, 123-125.
- Electrical, thresholds, cutaneous (rat), 624-627.
- Emmert's law, size-constancy and, 114-116.
- Emotion, fear, conditioning of, 69-74.
- Epilepsy, community, in the, 670 f.; convulsions, grand and petit mal, 670 f.

- Errata, 471, 655.
 Errors, pattern of, in delayed recall, 594-599.
 Estimations, loudness, of, 600-605.
 Extinction, resistance to, 123-125; training and resistance to, 451 f.
 Eye, electrical and photic stimulation of, 577-585; inter-, color differences, 386-394, 655.
 Eye-blink reflex, dark-adaptation and the 462 ff.
 Factor analysis, international conference on, 330 f.
 Fads and foibles, sociology, modern, in, 674 f.
 Familiarity, auditory, determinant of visual recognitive threshold, 634-636.
 Fantasy, analysis of, 158 f.
 Fear, conditioning of, 69-74.
 Field-theory; cortical integration, 528-540; stimulation, implications, 528-530.
 Figural inspection, influence of, an autokinetic illusion, 395-402.
 Figure-ground, auditory, 653 f.
 Films, research, 156.
 Fish. *see* Animal.
 Forgetting, memory-trace and, 132 f.
 Forms, effects of, in perception, 21-37, 38-48; circle and ellipse, memory for, 248-252.
 Freudiana, 669 f.
 Galvanic skin response, dark-adaptation and the, 462 ff.; semantic generalization and, 541-549.
 Gambling, discrimination in, 194-202.
Ganzfeld, *see* Vision, field.
 Gerontology, aging in industry, 664 f.; mental disorders in later life, 156 f.; verbal materials for old Ss, 628-630.
 Gestalt, hypothesis of memory trace not confined, 248-252.
 Goddard, Henry Herbert, *see* Biography.
 Groups, study of, 668 f.
 Healing, concepts of, new, 487 f.
 Hollingworth, Harry Levi, *see* Biography.
 Hypnosis, handbook of, medical, 661 f.; nature of, 661 f.; role in psychoneurotic and psychosomatic disorders, 661 f.; therapeutic applications, 661 f.
 Hypoxia, mental impairment in, 243-247.
 Illusions, autokinetic, influence of figural inspection on, 395-402; reversal perspective, 131 f.
 Imagination, projective testing and, 158 f.
 Individual, biochemical, the, 672 f.; differences, olfactory thresholds, 227-232; psychology (Adler), 151 f.
 Industrial psychology, social foundations, 159 f.
 Industry, aging in, 664 f.
 Inhibition, proactive and retroactive, 117-119; reactive, components of, 616-619; retroactive, extinction and, 123-125.
 Instructions, motivating, effect of, on learning, 620-623.
 Intelligence, United States, in, 668; Quotient (IQ), group, underestimations, 467.
 Intensity, odor, 471; factors associated with, 2 f.; smell, 1-20; visual, discrimination of, 211-218.
 Inter-American Congress of Psychology, *see* Meetings.
 Isomorphism, perception (visual) and brain stimulation, 528-540.
 Judgment: Adaptation-level, effect of, on, 347, 360.
 Assimilation, a context effect in, 352 f.
 Central tendencies of, 560-568.
 Clinical materials, of, 347-360.
 Confidence in, related to correctness, 432-434; scale of, 432-436.
 Context, effects of, in, 356 ff.
 Curvature, of, cues involved in, 361-375.
 Depth, of, 611-615.
 Line-lengths, 81-86.
 Loudness, 600-615; successive comparisons of, 117-119.
 Prediction of collisions, 110-113.
 Psychophysical, lifted-weights, 268-271; visual size (circles), 272-275; (triangles), 276-279.
 Shifts in, 351 f.
 Size, apparent, 114-116; children's vs. adult's, 106-110; circular areas, 560-568.
 Training, effect of, 434-436.
 Velocity, real and apparent, 102-105.
 Kinesthetic stimulation, motor performance and, 616-619.
 Language, foreign, use of, by various sciences, 314-316; speech, without, 541 ff.; thought, reality, and, 152 f.
 Law of effect, recent studies of, 642-648.
 Learning: Associative, role of repetition in, 186-193.

Learning (*continued*):

- Changes in, due to shifts in reinforcement ratios, 194-202.
- Compensatory, color weakness and, 126-127.
- Conceptual, 263-267.
- Conditioning, eyelid reflex and GSR, during various luminance levels, 462 ff.; of fear, 69-74.
- Discriminative, fish, 572 ff.
- Extinction, resistance to, and degree of training, 451 f.
- Fittingness, facilitating effect of, 450.
- Incidental, function of meaningfulness, 465 f.; intentional and, 253-257, 258-262; law and effect and, 642-645; nonsense figures, 447-450; verbal, 421-426.
- Insightful, 191 f.
- Intent, effect of, on 447 ff.
- Intentional vs. incidental, 465 f.
- Isolated vs. massed (Restorff effect), 258-262.
- Maze, Planarian, 128; variable pattern, 306 f.
- Motor, tracking, 49 ff., 75 ff.; transfer and retention of, 75-80.
- Probability-, 161-173, 586-593; concepts of, 162 ff.; multiple choices in, 174-185.
- Reward, cessation of electrical shock and, 421-426; effect of, on incidental learning, 642-648; punishment and, effect of, 219-226.
- Rote, effect of sensory deprivation on, 637-639.
- Semantic differential, a measure of meaning, 164 ff.
- Serial, associative value and instructions, effect of, on, 620-623; errors in delayed recall, 594-597; position-curve, 258-262.
- Theory, Spence's, critical review of, 141 ff.
- Training, intersensory transfer of, 283-285.
- Verbal, effect of instructions on, 92-96; materials for old and middle aged Ss, 628-630; patterns of response, 594-597.
- Limen, *see* Psychophysics: Limen; Threshold.
- Loudness, estimations of, by untrained Os, 600-605; scale of, 600-605. successive comparisons of, 117-119.
- Masking, odors, 3.
- Meaning, measurement of, 161-173; nationalism, of, 154; semantic differentiation and, 586-593.
- Measurement, interests of, vocational, 325 ff.; mind and matter, of, 329 f.
- Means, index of, vs. mean of indices, 467 f.
- Meetings: Eastern Psychological Association, 468 f.
- Inter-American Congress of Psychology, 470 f.
- Mid-Western Psychological Association, 469 f.
- Society of Experimental Psychologists, 320.
- Southeastern Psychological Association, 321.
- Western Psychological Association, 654.
- Memory, circle and ellipse, for 462 ff.; incidental and intentional, 253-257, 258-262, inhibition, proactive and retroactive, 117-119; lifted-weights, 253-257; trace, autonomous changes in, 132 f. visual figures, 248-252.
- Mental, deficiency; impairment in hypoxia, 243-247; *see also* Deficiency; Mental.
- Method: Constant-sum, in scaling, 81-86, 268-271.
- Identical stimuli, 248-252.
- Lifted-weights, memory for, 252-257.
- Scientific, in psychology, 155.
- Single stimuli, restriction of judgments in, 272-275.
- Thematic apperception technique, 158 f.
- Midwestern Psychological Association, *see* Meetings.
- Monkey, *see* Animal.
- Motion: Collisions predicted, 110-113.
- Perspective, 129-131.
- Translatory, apparatus for, 103, 291-294; explanation of, 102-105.
- Visual, perception of, during fixation and pursuit, 64-68.
- Motivation, effect of, on learning, 92-96; symposium on (Nebraska), 147 ff.
- Motor, reactions, visual, 56-63; skills, multipart, 550-559; tracking, analysis of its components, 49-55; transfer and retention after practice, 75-80.
- Movement, Aubert-Fleischel paradox, 64 f.; autokinetic effect, auditory, 133 f., illusion, 395-402; visual, apparent vs. physical, 102-105.
- Myopia: 'empty-field,' 376-385; high altitudes, at, 376-385.
- Nationalism, meaning of, 154.
- Nearsightedness, *see* Myopia.
- Necrology, *see* Biography.

- Neural quantum, hypothesis, 512-527.
 validity of, 526 f.; response-sequences and, 512-527.
- Nonsense-syllables, associative value of, effect of, on learning, 620-623.
- Obituary, *see* Biography.
- Occupations, psychology of, 663.
- Olfactory, *see* Apparatus, Olfactometer; Smell.
- Opinions, personality and, 665 f.
- Organization, cerebral cortex, of the, 673 f.
- Orientation, phenomenal, effect on form, 493-511.
- Perception: Attention and, 38-48.
 Autokinetic illusion, 385-402.
 Characteristics, 586-593.
 Color, inter-eye differences, 386-394.
 Concavity and convexity, 131 f.
 Contour, visual, 21 ff.
 Critical flicker frequency, under brain stimulation, 531 ff.
 Curvature, 361-375.
 Depth, 131 f.; textural gradients and, 611-615.
 Design of research in, 480-485.
 Emmert's law, 114-116.
 Fog, 403-410.
 Form, effects of, in, 38-48; effect of orientation, retinal and phenomenal, on, 493-511; threshold of, under brain stimulation, 534 ff.
 Letters, of, 40-48.
 Line-lengths, 81-86.
 under brain-stimulation, 530 f.
 Motion, visual, 110-113.
 Perspective, reversible, 535 ff.
 Redintegrative, of words, 97-101.
 Retinal fixation and, 38-48.
 Set, effects of, in, 38-48.
 Size, anxiety and discrimination of, 286-288; apparent, children vs. adults, 106-109; circles, 272-275; constancy, 114-116; triangles, 276-279.
 Targets, small aerial, 376-385.
 Velocity, 64-68.
 Visual, contour, 21; effect of, on brain stimulation, 528-540; field, of entire, 403-410; Gestalt principles, role of, in, 21 ff.; motion, apparatus for, 291-294, during fixation and pursuit, 64-68; organization, modes of, 30 ff.; phenomenal overlapping, determinants of, 21-27; space, 28 ff.; transfer of auditory to, 283-285.
- Perception (*continued*):
 Weight, of, 268-271.
 Words, tachistoscopically exposed, 38-41, 97-101.
- Performances, skilled, multipart, 550-559.
- Personality: Children, young, 327 ff.
 Clinical psychology and, Adler's, 151.
 Congenitally deaf, of the, 663 f.
 Opinions and, 665 f.
 Studies, clinical, of, 145 ff.
 Study of, 158 f.
 Theory, perspectives in, 666 f.
- Phenomenal, orientation, perception of form and, 493-511; overlapping, determinants of, 21-37.
- Pigeon, *see* Animal.
- Pitch, Tartini, 203-210, 646-649.
- Planaria, *see* Animal.
- Portrait, Terman, L. M., facing 333; acknowledgement of, 471.
- Practice, mass, on pursuit-rotor, 616-619.
- Preferences, monkey's, food, 87; objects not food, 87-91.
- Probability, words, occurrence of, 311-314; Ziff's law of, 311-314.
- Problem-solving: Anxiety and ability in, 445.
 Fluency and ability in, 445.
 Hierarchies of perceptual responses in, 442-446.
 Set, persistency of, in, 427-431.
 Sex differences in, 445.
 Subliminal stimulation in, 437-441.
 Verbal vs. manipulatory solutions, 289 f.
 Water jar test, 427-431.
- Prophecy, when it fails, 671 f.
- Psychology, annual review of, 323 ff.
- Psychophysics: Constant-sum method in scaling, 268-271; modified, 81-86.
 Central-tendency, effect of, 560-568.
 Judgments, effect of contrast and assimilation on, 560-568.
 Lifted-weights, point-assignment procedure, 268-271.
 Limen, absolute, 280 f.; relative, 282.
 Method of limits, 471; frequency procedure in, 280-282.
 Point of subjective equality in judgments of size, 563 ff.
 Principles of, applied to schizophrenic test-responses, 347-360.
 Single stimuli, judgments restricted, 272-275.
 Time-error in, 560-568.
- Rat, *see* Animal.

- Reaction-time, association, free, 606-610; discriminative and simple, 56-63; rate between successive, 56-63; rhythmic disturbances and, 56-63.
- Reading, backwardness in, 675 *f.*; difficulties, diagnosis and correction, 332.
- Reasoning, abstract, no sex difference, 446; impairment of, in hypoxia, 243-247.
- Recall, delayed, pattern of errors in, 594-599.
- Recognition: Attention and, 38-48.
- Cognition and, mathematical theory of, 322 *f.*
- Differential, word-length as a factor of, 316-318.
- Form, of, retinally and phenomenally oriented, 493-511.
- Retinal fixation and, 38-48.
- Span of, for letters and words, 38-48, 92-101.
- Threshold, auditory familiarity and, 634-636.
- Words, English and Hebrew, tachistoscopically presented, 38-48.
- Redintegration, perception of words and, 97-101.
- Reflex, eye-blink and dark-adaptation, 462 *ff.*
- Repetition, effect of, in learning, 186-193, 217-226
- Research, design, 480 *f.*, 658; films, 156; psychological, 658 *f.*
- Retention, autonomous changes in, 132 *f.*; perceptual organization facilitates, 447 *ff.*; tests of effects of delaying, 594-597.
- Retinal, fixation, perception, a function of, 38-48; orientation, perception of forms, and, 493-511.
- Review, annual, of psychology, 323 *ff.*
- Rewards, effect of, in learning, 219-226, 642-648.
- Rhythmic disturbances, effect of, on reaction-time, 56-68.
- Scale, confidence, 432-436; loudness, 600-605.
- Scaling, constant-sum method in, 268-271; psychological, line-lengths, 81-86.
- Science, methods in psychology, 155; philosophy of, 658.
- Sciences, foreign languages, use of, by different, 314-316.
- Scientific communication, strangled by reasons of military security, 130.
- Semantic, differential, 164 *ff.*; differentiation, word-association and, 586-593; generalization, 172 *f.*, conditioning and, 541-549.
- Sensory deprivation, rote learning and, 637-639.
- Set, effects of, in perception, 38-48; persistency of, in problem-solving, 427-431.
- Sex, difference, abstract reasoning, 446.
- Shape, influence of, on visual intensity, 211-218.
- Shock, expectancy of, GSR and, 541-549.
- Size: Apparent, function of distance, 106-109.
- Circular area, judgment of, 560-568.
- Constancy, 107 *f.*; effect of time on, 276-279.
- Discrimination of, anxiety and, 286-288.
- Emmert's law and, 114-116.
- Influence of, on visual intensity, 211-218.
- Skill, motor, analysis of its components, 49-55; transfer and retention of, 75-80.
- Sleep, effect of, on retention, 132 *f.*
- Smell, Adaption, 1-20, 471.
- Apparatus, odorometer, 1; olfactometer, 1.
- Intensity, 1-20; method of measuring, 3 *ff.*
- Olfacties, 1.
- Thresholds, differences in, 227-232.
- Social psychology, industrial and, 159 *f.*; personal influence, 157 *f.*; study of groups, 668 *f.*
- Society of Experimental Psychologists, *see* Meetings.
- Sociology, American families, 676 *f.*; fads and foibles in modern, 674 *f.*
- Southeastern Psychological Association, *see* Meetings
- Space, visual, 21-37.
- Stimulation, brain, polarized currents, with, 528-540.
- Subjective tones, Tartini and beat pitches, 203-210; 646-649.
- Suggestion, hypnotic, 661 *f.*
- Superstition, pigeons, in, 308-311.
- Tachistoscope, *see* Apparatus.
- Tartini, pitches 203-210; tones, 646-649.
- Teaching, college, 677.
- Terman, Lewis M., *see* Biography.
- Theory, assimilative and contrast, of central tendency, 560-568.
- Thinking, animal, 485 *f.*; problem, water jar test, 427-431.

- see also* Problem-Solving.
- Thought, language, reality, and, 152 f.; problem-solving, 289 f.
- Threshold: Critical flicker frequency, 551 ff.; electrical stimulation of the brain and, 531 ff.
- Cutaneous, electrical (rat), 624-627.
- Data, analysis of, 238-242; non-randomness of, 238-242.
- Differential, visual, 211-218.
- Form, absolute, 534 ff.
- Recognition, smell, 227-232.
- Recognitive, visual auditory familiarity a determinant of, 634-636.
- Smell, 2.
- Vision, luminance, 238-242.
- see also* Psychophysics, Limen.
- Time-error, theories, adaptation-level, 560-568; assimilative, 560-568; contrast, 560-568.
- Tones, beats, 210; difference, 203-210; subjective, 203-210, 646-649; summation, existence denied, 203, 210; Tartini pitches, 203-210, 646-649.
- Tracking, analysis of components, 49-55; pursuit, input partially controlled, 631-633, visual and kinesthetic stimulation and, 616-619; transfer and retention after practice, 75-80; visual targets, 110-113.
- Training, extinction, resistance to, and degree of, 451 f.; transfer of, intersensory, 283-285.
- Velocity, phenomenal, determinants of, 101-105.
- Vestibular, responses, electrically aroused, 530.
- Visibility, high altitudes, at, 376-385.
- Vision: Beats, 577-585.
- Binocular vs. monocular, in judgments of curvature, 367-369.
- Color perception, inter-eye differences, 386-394.
- Field of, at high altitudes, 376-385; entire, characteristics of, 403-410.
- Forms, memory for, 248-252.
- Intensity, discrimination, 211-218.
- Interocular color-effect, 318 f.
- Motion, apparent vs. physical, 102-105.
- Myopia at high altitudes, 376-385.
- Phosphenes, electrical, 577-585.
- Size, apparent, function of distance, 106-109.
- Stimulation, effect of, on motor performance, 616-619.
- Threshold, luminance, 238-242; recognitive, 634-636.
- see* Perception, Visual.
- Warmth, paradoxical, 640 f.
- Western Psychological Association, *see* Meetings.
- Word, association, free, 120-122; auditory vs. visual presentation, 233-237; length, differential recognition and, 316-318; random placement of spaces and letters, 311-314; recognition of, 38-48, 97-101.





An Outstanding New Text from Knopf

Clinical and Counseling Psychology



by **JOHN M. HADLEY**, *Purdue University*

This introductory survey is a comprehensive treatment both of therapy—the theories, aims, and techniques of counseling—and of the tools of diagnosis and evaluation—the means whereby the psychologist gets the information he needs to understand a client's current behavior and to plan intelligently for its modification. Each chapter has an extensive list of references cued into the text itself. The Glossary is full, precise, and clear.

Dr. Hadley is Director of Graduate Training in Clinical Psychology and Clinical Services at Purdue University.

6¼ x 9¾ 682 pp. \$6.75 text

ALFRED A. KNOPE, *Publisher*

501 Madison Avenue

College Department

New York



Subscriptions to

THE AMERICAN JOURNAL OF PSYCHOLOGY

\$7.00 per year—United States

\$7.50 per year—Foreign

Special Subscription Price to Members
of American Psychological Association



New subscriptions, renewals, and changes of
address should be sent to:

THE AMERICAN JOURNAL OF PSYCHOLOGY

Mezes Hall 209, The University of Texas
AUSTIN, TEXAS

Portraits of PSYCHOLOGISTS

Offprints of the portraits of Psychologists that have been published in The American Journal of Psychology are now available. The series includes the following:

G. Stanley Hall
Edmund C. Sanford
James Ward
Franz Brentano
Edward B. Titchener
Margaret Floy Washburn
Herman Ebbinghaus
Wilhelm Wundt
Howard C. Warren
G. E. Müller
Carl Stumpf
Kurt Koffka
William James
Raymond Dodge
John A. McGeech
James McK. Cattell
Max Wertheimer
Joseph Jastrow
Charles S. Myers
Pierre Janet
James R. Angell
Knight Dunlap
Edward L. Thorndike
Carl E. Seashore
Oswald Külpe
Clark L. Hull
David Katz
Walter S. Hunter
Harvey A. Carr
Madison Bentley
Louis L. Thurstone
Robert M. Yerkes
Lewis M. Terman

Separates—25¢ each

Set of 33—\$7.00

**THE AMERICAN JOURNAL
OF PSYCHOLOGY**

Mezes Hall 209, The University
of Texas, Austin, Texas

MORE POWER TO YOU

A Teen Guide for Self- Understanding



by John and

Doratheia

Crawford



Written just
for teen-agers,
in their
language,

MORE POWER TO YOU is packed full of ideas on how to gain clearer self-understanding, self-acceptance and self-control. The lively style of the book is made even more appealing to youthful readers by the use of diagrams, tables, illustrations and self-quizzes.

Real-life situations at the end of each chapter present "what-would-you-do" problems for discussion with each other or with parents.

\$2.50 at all bookstores.

MUHLBERG PRESS
Philadelphia



